이학박사학위논문

대안적 텔로미어 유지기작의 새로운 DNA 복제단위 연구

Studies on novel DNA signatures of telomerase-independent telomere maintenance in *Caenorhabditis elegans*

2016년 2월

서울대학교 대학원

생명과학부

김 천 아

ABSTRACT

Studies on novel DNA signatures of telomerase-independent telomere maintenance in *Caenorhabditis elegans*

Chuna Kim

Dept. of Biological Sciences

The Graduate School

Seoul National University

Telomeres at linear chromosomal ends cannot be replicated by DNA polymerase. This is called the end replication problem. If this problem is not solved, chromosomes will gradually be shortened along cell divisions. In order to maintain chromosome integrity through generations, organisms have invented telomere maintenance mechanisms. A specialized reverse transcriptase, telomerase, has evolved in ancient eukaryotic cells. Interestingly, telomerase has been lost a few times during evolution. In this situation, telomerase-independent mechanisms were adopted. Among telomerase-independent mechanisms, alternative lengthening of telomeres (ALT) has been studied in cancer formation. However, many studies on ALT did not have organismal perspectives. Here, I establish a model for organismal telomere maintenance after telomerase loss. Using *Caenorhabditis elegans* telomerase mutant,

I have generated telomerase-independent survivors. These survivors mobilized specific internal sequence blocks for telomere lengthening, which we named TALTs (Templates for ALT). These TALTs region consist of a block of genomic DNA flanked by telomere-like sequences. Telomerase-independent survivors *trans*-duplicate TALTs, which is already *cis*-duplicated to chromosome ends, across the telomeres of all chromosomes. The switch from the telomerase-mediated to the TALT-mediated telomere maintenance mechanism shows a snapshot of telomere evolution. In addition, the molecular event of the *cis*-duplication of TALTs suggests a novel mechanism for subtelomere formation

---

Keywords: *C. elegans*, genetics, telomere, telomere maintenance mechanism, alternative lengthening of telomeres (ALT), template for ALT (TALT), animal model for telomerase-independent survivor

Student Number: 2009-20328

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

1. Telomeres: protecting chromosomes ends

Without exception, eukaryotic nuclear genomes are contained in linear chromosomes. The ends of linear chromosomes are called ‘telomeres’ (Blackburn, 1991), after the Greek ‘telo,’ meaning ‘end’, and ‘mere,’ meaning ‘part’. Telomeres were first proposed in studies of Drosophila by Hermann Joseph Muller (Muller and Herskowitz, 1954). According to his description, most regions of Drosophila chromosomes are severely rearranged after irradiation, whereas the terminal regions of the chromosomes are not affected. Thus, he proposed that chromosome ends might be protected by specialized structures. Later, Barbara McClintock found that broken chromosomes frequently fuse with other chromosomes in irradiated maize (McClintock, 1942). Fused chromosomes induce breakage-fusion-bridge (BFB) cycles, which in turn cause severe chromosome losses at the fusion sites during mitosis. Therefore, McClintock proposed that normal chromosome must be distinguished from broken chromosomes in order for cells to maintain chromosome integrity.

From the standpoint of DNA synthesis, the chromosomal ends pose another problem. After the molecular details of DNA structure and replication machinery were revealed, Olovnikov recognized that the end of a lagging strand could not be replicated by conventional DNA polymerase (Olovnikov, 1973). Unless this problem were solved, chromosomes would gradually become shorter. Olovnikov called this problem the ‘end-replication problem’. Thus, some other type of DNA replication is clearly responsible for maintaining chromosome integrity.

2. Telomerase and alternative lengthening of telomeres

Eukaryotes have to solve two problems associated with linear chromosomes; the ‘end-protection problem’ suggested by McClintock, and the ‘end-replication problem’ suggested by Olovnikov. Living organisms use diverse strategies to solve the end-protection problem, including a covalently closed hairpin structure and a higher-ordered nucleoprotein structure (Fulcher et al.,

2014). The most well-known example is the shelterin complex, which is a telomere binding protein complex that form a T-loop structure in mammals. The end-replication problem is solved by various mechanisms such as recombination, retro-transposition and reverse-transcription to add sequences to the ends of chromosomes (Teng and Zakian, 1999). Many eukaryotic cells including humans use specialized reverse transcriptase called telomerase (Greider and Blackburn, 1985).

Not all cells can consistently avoid the end-replication problem. For example, in humans, telomerase is expressed only in germ cells and stem cells (Shay and Wright, 2010). The telomeres of somatic cells gradually shorten. If telomeres shrink below a certain threshold, cells enter replicative senescence and stop dividing (Campisi, 2001). From another point of view, replicative senescence can be considered as a protective strategy allowing somatic cells to avoid the catastrophic destruction of chromosomes and cell death caused by critically shortened chromosomes. An exception to this situation occurs in cancer cells, which can divide indefinitely without losing telomeres. To achieve unlimited

proliferation, cancer cells must overcome the replicative senescence resulting from the lack of telomerase activity. Thus, most cancer cells re-express telomerase (Shay and Bacchetti, 1997).

However, approximately 15% of cancer cells do not show telomerase activity, and this mechanism of telomere maintenance is called alternative lengthening of telomeres (ALT) (Bryan et al., 1997; Cesare and Reddel, 2010). By definition, ALT is a telomerase-independent telomere lengthening mechanism, and the "gold-standard" of ALT is the lack of telomerase activity. Some types of cancer show predominance of ALT (Heaphy et al., 2011b). In particular, sarcomas of mesenchymal origin, including osteosarcoma, liposarcoma (Henson et al., 2005), glioblastoma multi-forme (GBM) (Schwartzentruber et al., 2012), and pancreatic neuroendocrine tumors (PanNET) (Heaphy et al., 2011a), represent clear cases of ALT. ALT can be activated in telomerase deficient condition, and may therefore serve as back-up mechanism in telomerase positive cells. However, the exact mechanism by which ALT can lengthen telomeres is not yet well understood.

3. Ancient mechanisms of telomere maintenance

Although the widespread utilization of telomerase throughout phylogenic tree indicates that telomerase emerged early in eukaryotic evolution (Malik et al., 2000), linear chromosomes and telomerase might not necessarily have evolved simultaneously (Garavís et al., 2013). Many hypotheses have been proposed to explain how early eukaryotes protected the ends of linear chromosomes during the gap period in which telomerase did not exist (Curcio and Belfort, 2007; De Lange, 2004; Nosek et al., 2006).

One conceivable explanation is that newly evolved eukaryotic cells might have used a mechanism that was already present in prokaryotes. Mechanisms involving hairpin structures, tandem repeats, retrotransposition, and recombination might be ancient forms of chromosome maintenance (Fulcher et al., 2014). These ancient mechanisms can be used again for telomere maintenance after telomerase loss. In yeast, some cases of telomerase defective strains use retrotransposon for telomere maintenance (Moore and Haber, 1996; Teng et al.,

1996). In addition, there are interesting cases in nature in where telomerase-independent telomere maintenance mechanisms (TMMs) act (Mason et al., 2011).

These mechanisms may guide the understanding of ALT in cancer cells, because ALT in cancers resembles telomerase-independent TMM found in nature. Therefore, it would be appropriate to extrapolate what happened to organisms that lost telomerase during evolution to understanding the mechanism of ALT in cancer.

4. Alternative means of lengthening telomeres in nature

During the evolution of animals and plants, telomerase has been independently lost several times (Mason et al., 2011). The loss of telomerase must have first caused a crisis at the chromosomal end sequences. Many pieces of evolutionary evidence show that telomeres were replaced with a tandem array of DNA elements instead of simple telomeric repeats, which are added by telomerase. The most well-studied case invovles the order Diptera, which lost telomerase

more than five times as early as 260 million years ago (Mason et al., 2015). Drosophila is a representative case (Garavís et al., 2013). Drosophila use the end-specific retrotransposons HeT-A, TART and TAHRE (collectively called HTT), for telomere maintenance. In addition, the telomeres of *Chironomus tentans*, another Dipteran, show complex tandem-repeat patterns. The telomeres of many other insects and plants have tandem-repeat structures, including transposon-like or rDNA sequences (Dvořáčková et al., 2015). In plants, some Solanaceae subgroups do not have telomerase activity. The Cestrum subgroup of Solanaceae has satellite repeat sequences at the telomeres (Sykorova et al., 2003). Alliaceae species also have unusual telomere sequences. The telomere of *Allium cepa* contain satellite sequence and rDNA repeats (Pich and Schubert, 1998).

5. Purposes of this study

As described above, in cases of defective telomerase activity or gene loss, sequences other than the canonical telomeric repeats have frequently been used as

alternatives in telomere maintenance. Because ALT studies have primarily been focused on cancer in mammals, ALT is typically viewed as a problem for human health rather than a natural adaptive mechanism for survival. In addition, most cellular studies considered cellular rather than the organismal perspective. In this research, I have concentrated on exploring the possibility that ALT provides a selective advantage at the organismic level after telomerase loss.

To investigate the mechanism underlying ALT activation, I used a genetic model of the nematode *Caenorhabditis elegans* in which the telomerase gene *trt-1* is deleted and telomere-mediated sterility occurs after 14–18 generations (Meier et al., 2006). This model system enables me to investigate the mechanism of ALT by genetically preventing telomerase activation because survivors that are maintained beyond the expected number of generations can arise only through a telomerase-independent ALT mechanism.

Previously, two different approaches have been taken to identify ALT survivors in *C. elegans.* One is based on maintaining telomerase mutants by large-scale transfer of animals at each generation, in which very rare natural survivors

can arise (Cheng et al., 2012). The exact mechanism of escaping sterility in these natural survivors has not been elucidated. The other is candidate approach, which includes double mutant of telomerase with single-strand telomere binding proteins, *trt-1;pot-1* and *trt-1;pot-2* (Lackner et al., 2012; Shtessel et al., 2013). These survivors show a heterogeneous telomere length pattern that is similar to the ALT pattern in cancer, but stable maintenance of survivors has not always been successful by small-scale transfer (Shtessel et al., 2013). From these studies, I thought that the instability of telomerase defective *C. elegans* survivors exhibit characteristics of ALT cancer. Therefore, in this research, I tried to find telomerase-independent survivors that can stably be maintained and investigate on the changes in telomere sequences.

I found that the stably maintained survivor lines utilize an internal genomic region as a template for telomere lengthening. The amplified region in ALT was named TALT (Templates for ALT). I defined two TALT regions, both of which share a common sequence structure consisting of a block of genomic DNA flanked by telomere-like sequences. ALT survivors utilized, depending on their

genetic backgrounds, either one of two specific internal genomic regions as templates for telomere lengthening. These regions have *cis*-duplicated to the subtelomeric region of the same chromosome, forming reservoirs that are incorporated into all telomeres on ALT activation. The TALTs represent a novel DNA feature of ALT, whereby internal genomic regions are utilized to protect chromosome ends in the absence of telomerase. TALT amplification will provide evolutionary perspective of subtelomere formation.

Materials and Methods

***C. elegans* strains and culture.** Worms were cultured at 20°C under standard culture conditions (Brenner, 1974). The following strains were used in this study: Bristol N2 wild strain, Hawaiian CB4856 wild isolate, *trt-1(ok410)* I (Meier et al., 2006). N2 *trt-1(ok410)* was outcrossed with Hawaiian CB4856 wild isolate to produce CB4856 *trt-1(ok410)*. To ensure the ALT was activated in outcrossed progeny, F2 worms were grown at least 20 generations and worms with low fecundity were excluded. To maximize the outcrossing effect, SNVs of all chromosome markers were checked. The *trt-1(ok410)* mutation was confirmed by PCR and WGS. The *trt-1(ok410)* strain was outcrossed with N2 wild type to produce an early generation of N2 *trt-1(ok410)*.

EMS treatment. Synchronized L4 worms were treated with 50 mM EMS in M9 buffer for 4 hours. After 4 hours of recovery, treated P0s were allowed to lay eggs for 12 hours. F1 worms were isolated by removing P0 worms. Initially ~100 F1 eggs were transferred to fresh plate, then from the F2 generation onwards, 10~15 worms were transferred manually at every generation.

Feeding RNA interference. *E. coli* HT115 expressing dsRNA were grown in LB with 1 mM ampicillin at 37°C overnight and seeded on to NGM plates containing 1 mM IPTG and 1 mM ampicillin. At every generation, 10-15 L1 larvae were transferred to fresh RNAi media plates.

Telomere florescent *in situ* hybridization (FISH). As cells are highly dividing in embryo stage in *C. elegans*, it is most plausible condition that ALT is activated. Eggs were isolated by bleaching adult worms. Eggs were fixed in 2% paraformaldehyde (PFA). Tubes containing eggs were frozen in liquid nitrogen and thawed in warm water twice in order to crack the eggs. Eggs were settled on a slide coated with poly-lysine. The slide was washed 3 times with phosphate-buffered saline containing 0.1% Tween-20 (PBST) to remove residual PFA. The slide was incubated in acetone and methanol for 5 minutes each at -20°C and was then rehydrated in 2X SSC (0.3 M NaCl, 0.03 M sodium citrate) containing 0.1% Tween-20. The slide was blocked for 1 hour with prehybridization solution (3X SSC, 50% formamide, 10% dextran sulfate, 50 μg/mL heparin, 100 μg/mL yeast

tRNA, 100 μg/mL salmon sperm DNA) at 37°C. PNA-$(TTAGGC)_3$ probe was hybridized for 16 hours in humid chamber at 37°C. Slides were washed twice in wash buffer (2X SSC and 50% formamide) for 15 minutes at 37°C. After washing 3 times with PBS-T, slides were counter-stained with DAPI and mounted with anti-bleaching solution Vectashield (Vector Laboratory). The samples were imaged using a confocal microscope (LSM200, Zeiss). TALT1 probes were labeled with digoxigenin (DIG) further visualized by FISH. After FISH, slides were blocked with PBS-T containing 5% BSA for 1 hour at room temperature. Slides were stained with rhodamine conjugated anti-DIG antibody for 3 hours. After washing in PBS-T twice, slides were mounted and observed as described above. The telomere signal was quantified using TFL-TELO software (Dr. Peter Lansdorp, Terry Fox Laboratory, Vancouver).

Telomere southern blot. For genomic DNA preparation, worms were harvested and washed 5 times in M9 buffer. Worms were lysed in lysis buffer for 8 hours (100 μg/mL proteinase K, 50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM $MgCl_2$,

0.45% NP-40, 0.45% Tween-20, 1% beta-mercaptoethanol) DNA was extracted using phenol:chloroform extraction and ethanol precipitaion. DNA in TE buffer was treated with RNAse (10 μg/mL) for 2 hours and re-extracted, before being dissolved in TE buffer. For Southern hybridization, 5 μg of DNA was treated with 1 unit of restriction enzyme and then separated by gel electrophoresis either using standard equipment or Pulsed Field Gel Electrophoresis equipment. Gels were blotted by capillary transfer on to the Zeta probe membrane (Bio-Rad) overnight. The membranes were cross-linked using a UV cross-linker and hybridized with the Southern probe in DIG Easy Hybridization buffer at 42°C for 16 hours. The membrane was then washed twice at room temperature in 2X SSC, 0.1% SDS and twice at 42°C 0.2X SSC, 0.1% SDS. The DIG-labeled probe was detected on an ImageQuant LAS-4000 biomolecular imager (GE healthcare) using an anti-DIG-AP antibody chemiluminescence detection kit (Roche). The $(TTAGGC)_{30}$ probe was labeled with DIG-UTP by PCR-amplifying telomere sequences cloned in a T-easy vector. Probes for TALT1 and TALT2 were labeled with DIG-UTP using primers targeting unique region in TALTs.

C-circle assay. Worm genomic DNA was digested with TALT1 non-cutting restriction enzyme mix (*Nhe*I, *Bam*HI, *Dra*I, *Apa*I, *Nde*I, *Xho*I, *Nco*I, *Sac*I 4 units/μg each). Restricted DNA was purified by Phase Lock Gel (5 PRIME) and ethanol precipitation. 10 μL of sample was mixed with 10 μL 0.2 mg/mL BSA, 0.1% Tween-20, 1 mM each dATP, dGTP, dTTP, dCTP, 1X phi29 buffer with or without 7.5 units phi polymerase (NEB) and incubated at 30°C for 8 hours then at 65°C (polymerase inactivation) for 20 minutes. For dot blotting, sample were diluted with 60 μL 2X SSC and blotted onto nylon membrane. DNA was crosslinked with UV onto the membrane and hybridized at the 62°C with $(GCCTAA)_4$-digoxigenin probe in DIG easy hybridization buffer (Roche). C-circle amplified signal was detected by DIG detection kit (Roche) according to the manufacturers' instructions. *pot-1(tm1620)* and *pot-2(tm1400)* were used as positive controls as they are reported to contain elevated levels of C-circles compared to N2 wild type (Lackner et al., 2012; Shtessel et al., 2013). For sample loading confirmation, we stripped by 0.2 M NaOH, 2% SDS and re-hybridization with $(GCCTAA)_4$-digoxigenin probe (denatured blot).

BAL 31 exonuclease treatment. 5 μg of DNA was treated with 10 units of BAL 31 at 30°C in 1X BAL 31 buffer. Reactions were stopped with the addition of EGTA (25 mM final concentration). DNA was collected by ethanol precipitation and digested with non-cutting restriction enzyme mix (*Nhe*I, *Bam*HI, *Dra*I, *Apa*I, *Nde*I, *Xho*I, *Nco*I, *Sac*I).

Sequence analysis of telomere-proximal telomere region junctions. To analyze the exact sequence of the junction of telomere and proximal telomere region, PCR products of junctions were sequenced. For PCR reactions, the forward primer was designed against telomere-adjacent DNA to elongate into the telomere and reverse primer was designed from TALT1 specific sequence (the first exon of T26H2.5). CS1 genomic DNA and N2 genomic DNA were used as template. 30 PCR cycles were performed with primers annealing at 60°C and elongation progressing for 3 minutes. After electrophoresis in 1% agarose gel, the CS1-specific amplicon was gel-extracted and sequenced. Primers used for sequencing were the same as those used for PCR reactions.

Whole genome sequencing. DNA was fragmented to 300 bp and sequencing libraries were constructed with an average insert size of 430 bp using standard Illumina protocols. Libraries were run on an Illumina Hiseq2000 sequencing platform.

Variant Discovery. To analyze variant induce by EMS and CB4856 associated variant, WS243 version of *C. elegans* reference genome and annotation data were acquired from Wormbase website (www.wormbase.org). For preprocessing the raw sequencing data, Trimmomatics was applied so that illumina adapter sequences were removed (Bolger et al., 2014). 3′ and 5′ ends of low quality reads were trimmed. Preprocessed reads were then aligned to *C. elegans* reference genome (WS243) with the Burrows-Wheeler Aligner software using default parameters (version 0.7.5a) (Li and Durbin, 2009). Before calling variants, Picard's SortSam and MarkDuplicates and GATK's indel realignment (McKenna et al., 2010) and base quality score recalibration (BQSR) was applied. For the indel realignment, very stringent parameters were used due to the probability of

considerable genetic difference between the reference assembly and survivor lines. From reference annotation (WS243), SNVs and indels sourced from the million mutation project were selected and used in BQSR. Variant discovery and genotyping were performed across all samples simultaneously using GATK's haplotype caller as per GATK Best Practices, then standard hard filtering was applied with minimum raw depth of coverage of 4.

TALT1 contig generation. Since TALT1 is amplified from CB4856 allele (TALT1 R) that is absent from genome assembly (based on the N2 genome), we reconstructed TALT1 R region through *de novo* assembly of CB4856 *trt-1* sequencing data using CLCworkbench7 (Qiagen) with default setting. Using BWA, reads of CS1 were collected if only one of the pair was mapped on the reconstructed TALT1 reservoir. Collected reads were used to generate contig sequences, which were then aligned to N2 genome using BLAST.

Whole genome sequencing analysis for telomere reads. WGS data was aligned to the ce10/WBcel215 reference assembly using bwa (Li and Durbin, 2009) and converted into bam files using Samtools (Li et al., 2009). Bam files from different worm strains were analyzed for reads predicted to derive from within telomeres using previously described software (http://sourceforge.net/projects/motifcounter/) (Conomos et al., 2012). Briefly, reads which contained at least 6 canonical telomere repeats (TTAGGC) were extracted and saved into separate files. The number of reads was counted and normalized to total read density to infer telomere length. The saved files were subsequently used to analyze split-pair regions and for variant calling.

TALT bioinformatic analyses. To identify regions that were associated with the telomere, files containing telomeric reads were used to extract their associated paired-ends. Reads in which both pairs were telomeric were excluded, leaving only reads in which one pair mapped to a genomic location. In addition to subtelomeric regions and loci flanking interstitial telomeres, the majority of the

reads mapped to TALT1 in the case of CS1 and CS2, and TALT2 in the case of NS1 and NS2. Average genomic coverage across the genome was calculated using genomecov from the BedTools package (Quinlan and Hall, 2010), after the removal of duplicates and low mapping quality (q<20) reads in Samtools. The read depth was assessed at every nucleotide in the genome and normalized by dividing by the average coverage to give a fold-change metric for read depth. The fold-change read depth from the parental strains was subtracted from the ALT survivors to account for strain-specific variations across the genome. Line plots of genome-wide and locus-specific read depth changes were created in R. Putative additional TALT regions were identified by analyzing the *C. elegans* assembly with the Homer scanMotifGenomeWide function (Heinz et al., 2010) for $(TTAGGC)_7$ with up to 3 mismatches to find all genomic ITS elements. The distance between each ITS was calculated, and data was extracted if this distance was between 100-2,000 bp. All candidate regions were manually confirmed, the 5′ and 3′ ITS length recorded, and any genes noted.

Telomere variant calling. Telomere variants were analyzed from files containing telomere reads. The number of each variant type was counted within this subset of reads from each library using the program motif_counter (Conomos et al., 2012). The presence of each variant repeat was assessed within this subset of reads, and their frequency was calculated and used to generate pie charts.

Measurement of TALT copy numbers. TALT copy numbers were measured from worm total genomic DNA by using quantitative real time PCR (qPCR). We calculated TALT copy number normalized by a single copy gene, *act-1*. Each reaction performed on the iCycler iQ5 (Bio-Rad) using following thermal profile: 95°C for 3 minutes; 40 cycles of 95°C for 20 seconds, 60°C anneal for 20 seconds, 72°C extend for 30 seconds along with 81 cycles of melting curve from 60°C to 95°C. The reaction components are as follows : 10 µL 2X SYBR Green mix (Bio-Rad), 1 µL each of 10 µM forward and reverse primers, 7 µL sterile water and 1 µL total genomic DNA (100 ng/µL) in 20 µL reaction

mRNA sequencing analysis. To align the preprocessed RNA-seq reads to reference genome and estimate the expression level of transcripts, we followed the Tuxedo pipeline (Tophat2 and Cufflinks pipeline) (Trapnell et al., 2012). First, Tophat2 mapped RNA-seq reads to the *C. elegans* genome using given transcript annotation (WS243) (Kim et al., 2013). Next, using cuffquant and cuffnorm software in Cufflinks package (Trapnell et al., 2010), we estimate the abundance of known transcripts in normalized RPKM across all samples.

Gene set enrichment analysis (GSEA). The gene ontology (GO) annotation data for *C. elegans* were downloaded from the Wormbase website (www.wormbase.org). Gene sets that change upon irradiation was obtained from ref. 10. For all pairs of ALT samples and a wild type, Gene Set Enrichment Analysis (GSEA) was performed using the gene set data obtained above (Subramanian et al., 2005).

Phylogeny construction. Population phylogenies were constructed by using a fasta file listing the genome-wide variant positions across all strains, and subsetting by regions as desired. Muscle (version v3.8.31) was used to generate a neighbor-joining tree. The ape (version 3.4) and phyloseq (version 1.12.2) R packages were used for plotting.

Wild isolates TALT analysis. The depth of coverage in putative TALT regions was calculated from bam files of sequenced wild isolates using a custom program (available at github.com/AndersenLab/bam-toolbox). Coverage within regions was normalized to average genome-wide coverage.

Results

1. Isolation of telomerase-independent survivors of *C. elegans*

To find telomerase-independent survivor, two natural isolates (N2, CB4856) of *C. elegans* telomerase mutant, *trt-1(ok410)* were treated by alkylating agent EMS (Figs. 1, 2). Both wild-type backgrounds show the most divergent number of single nucleotide variants (SNVs) (Andersen et al., 2012). Since no *trt-1* mutant CB4856 strains are available, *trt-1* CB4856 was constructed a strain bearing this mutation by extensive outcrossing of *trt-1* N2. In *trt-1(ok410)*, telomere-mediated sterility occurs after 14-18 generations. After EMS treating, survivors could be maintained by typical transfer which is moving 10-15 individuals every generation (Fig. 3). 6 independent survivors (named as CS1-CS6) from 80 EMS-treated plates of CB4856 *trt-1* and 5 independent survivors (named as NS1-NS5) from 200 EMS-treated plates of N2 *trt-1* were found. All 6 CS survivors and 2 of the NS survivors (NS1 and NS2) could be maintained with no gross phenotype for at least 300 generations by transferring 10 first larval stage (L1) worms at each generation (Seo et al., 2015) (Fig. 4a). The remaining three

NS survivors (NS3-NS5) were able to be maintained only by massive transfer of worm every generation (Fig. 4b). The chromosome number of all the telomerase-independent survivors decreased due to chromosome fusions, suggesting that telomeres were critically shortened before survivors were established (Fig. 5).

Similar to ALT in human cancer cells, the telomere lengths of survivors were longer than those in the starting strains when assessed by fluorescent in situ hybridization (FISH), terminal restriction fragment (TRF) analysis and whole-genome sequencing (WGS) analysis (Figs. 6, 7, 8). In addition, long telomere signal was degraded by BAL 31 exonuclease (Fig. 9). These results suggested that telomere was lengthened in survivors (Seo et al., 2015).

Unlike those of the starting strains, telomere sequence of these survivors was cut by a restriction enzyme that does not cut canonical telomere sequence (Figs. 10, 11). Restricted telomere sequence that were hybridized by telomere repeat DNA probe showed discrete banding patterns, which is distinct from smear patterns produced by canonical telomeres in TRF analysis. All 6 CS survivors had the same discrete banding patterns that resembled each other, and the 2 NS

survivors also had discrete patterns (Figs. 12, 13). However, the patterns from the NS survivors and CS survivors did not resemble each other (Fig. 14). Therefore, it was hypothesized that specific non-telomeric units might have been inserted into telomere and that the survivors from different genetic backgrounds use different non-telomeric units for telomere maintenance (Seo et al., 2015).

In summary, all 6 CS survivors and 2 of the NS survivors, which can be stably maintained by means of a typical transfer, have discrete banding patterns in TRF analysis. The remaining 3 NS survivors did not have discrete banding patterns but instead showed a smear telomere pattern, similar to that of wild type (Fig. 13). Only the NS3-NS5 survivors did not show discrete banding patterns, and were maintained only by massive transfer. Therefore, survivors containing non-telomeric units can be maintained stably (Fig. 4).

2. Telomere maintenance by TALT in survivors.

To find telomere-inserted unit, whole genome sequencing analysis of the survivors was performed. Using paired-end sequencing, certain reads having telomere repeats in either of the paired-end reads was collected and aligned to reference N2 genome (Fig. 15). Interestingly, many reads from both survivors were aligned to a single internal region in the genome. In N2 survivors, telomere-including reads were located in telomere-adjacent region in the left arm of chromosome I (Fig. 16). In CB4856 survivors, these reads were located in an internal region of chromosome V (Fig. 17). These amplified sequences were named TALT (Template for ALT) (Seo et al., 2015). Previously, Karlseder group showed *trt-1(ok410);pot-1(tm1620)* double mutant can lengthen trans-generational lifespan than *trt-1(ok410)* (Lackner et al., 2012). However, this mutant can neither be maintained indefinitely by small-scale transfer nor has amplified genomic region like TALT survivors (Figs. 18, 19). The CB4856-type TALT was named TALT1, and the N2-type TALT was named TALT2.

To identify whether TALTs are used for telomere lengthening of telomerase-independent survivors, a bioinformatics approach was tried. All pair-end sequence reads that mapped to TALT1 were extracted and used overlapping mate pairs to construct contiguous fragments ranging from 200 bp to 1.7 kb in length (Fig. 20). 5 contigs that spanned both TALT1 and specific chromosomal ends were generated (Table 1). This result indicates that TALT1 is located to chromosome ends. To confirm this result, PCR was performed using TALT1-specific primer with chromosomal end-specific primers. PCR amplicons were produced only from the CS survivor (Fig. 21 and Table 2). Also in NS survivors, PCR amplicons were made using the TALT2-specific primer with chromosomal ends-specific primers (Seo et al., 2015) (Fig. 22 and Table. 2).

Only in CS survivors did TALT1 probe perfectly co-localize with telomere repeat probe in FISH (Fig. 23). In addition, TALT2 probe perfectly co-localized with telomere probe only in NS survivors in FISH (Fig. 24). The long telomere of survivors was not cleaved after treating restriction enzyme that did not cut within TALT sequence (Figs. 25, 26). Finally, quantitative PCR (qPCR)

analysis was performed using TALT-specific primers. This experiments showed that only telomerase-independent survivors have higher TALT copy number (Fig. 27). The copy number of TALT1 specifically increased in CS survivors, while the copy number of TALT2 specifically increased in NS survivors (Seo et al., 2015) (Fig. 28).

Both TALTs share a particular feature: a unique sequence flanked by telomere repeats (Fig. 29). The TALT1-unique sequence includes the promoter and one exon of the T26H2.5 ORF, and TALT2-unique sequence includes an intergenic region. The expression of unique sequence is slightly increased in survivors compared to *trt-1* mutant (Fig. 30). Flanking telomere sequence have telomere variant repeats. Both survivors have different telomere variant repeats in their flanking sequence (Fig. 31). TALTs are probably duplicate as units to chromosome ends of the survivors in a repetitive head-to-tail patterns (Fig. 32).

3. ALT mechanism involves TALT duplication in *cis* and in *trans.*

TALT2 (N2) aligned to a telomere-adjacent region whereas TALT1 (CB4856) aligned to internal genomic region. Both N2 and CB4856 have extensive divergence in single nucleotide polymorphisms (SNPs) because of geological distance of their origin. TALT1 from CB4856 also have SNPs compared to TALT1 from N2. TALT1 has two SNPs in whole genome sequencing reads (Fig. 33). However, I could not find any SNP difference between both strains by Sanger sequencing of the internal region of chromosome V (Fig. 34). Interestingly, SNPs of TALT1 were located in telomere-adjacent region in the right arm of chromosome V (Fig. 34). Reason of TALT1 mis-alignment was structural variation of CB4856 in telomere-adjacent region. TALT1 region of CB4856 had distinct restriction patterns in Southern blot (Fig. 35). It was confirmed by PCR using TALT1-specific primer and chromosome end primer. Only in CB4856, PCR amplicon between right arm of chromosome V and TALT1 was made (Fig. 36).

In qPCR experiments, copy number of TALT1 showed 8-fold increase in CB4856 than that of N2 (Fig. 37). TALT1 showed a tandem repeat pattern in PCR (Fig. 38). To verify exact TALT1 loci, long-read whole genome sequencing experiment was performed by PacBio sequencer. Median read length of PacBio sequencing was about 10 kb. In this sequencing, it was confirmed that telomere-adjacent region of CB4856 have multiple copies of TALT1 (Fig. 39). Hence, TALT1 exists in two genomic regions, internal and telomere-adjacent regions. TALT2 also exists in two genomic regions, internal and telomere-adjacent region. It was confirmed by qPCR experiments and BLAST assays (Fig. 40). Interestingly, flanking sequences of TALT2 in internal and telomere-adjacent regions were different from each other (Fig. 41).

Among two genomic regions of TALTs, it was the telomere-adjacent region that was increased in survivors. CB4856-derived alleles of TALT1, located in telomere-adjacent region, were amplified in CS survivors (Fig. 33c). Also in NS survivors, telomere-adjacent variants of TALT2 were increased (Fig. 42). Both CB4856 and N2 have internal regions of TALT1 and TALT2. However,

telomere-adjacent TALT1 exist only in CB4856 whereas telomere-adjacent TALT2 exist only in N2 (Fig. 43). Therefore, strain-specific duplication from internal to telomere-adjacent region occurred during evolution. This mechanism was named *cis*-duplication. *Cis*-duplicated TALT can be recruited to all chromosome ends for telomere lengthening after telomerase loss. This mechanism was named *trans*-duplication (Fig. 44).

Cis-duplication had occurred before telomerase loss. In *C. elegans*, many wild isolates exist. *Cis*-duplication of TALT exist in multiple wild isolates (Figs. 45, 46). Strains that have experienced *cis*-duplication of TALTs do not share common ancestor. Therefore, TALTs have been independently duplicated in multiple wild isolates.

4. Inducer of TALT *trans*-duplication.

What is the inducer of TALT *trans*-duplication to telomeres? Studies from Ahmed lab could not find distinctive telomere band pattern in non-EMS treated survivors by chunking (Cheng et al., 2012). TALT duplication emerged only in the EMS-treated condition. Therefore, EMS would induce TALT *trans*-duplication. EMS is commonly used as an inducer of point mutation in genetic studies of *C. elegans*. Hence, ALT suppressor mutation that might be induced by EMS was first traced. SNP mapping approach was performed in both survivors. In CS survivors, unrelated mutations were eliminated by outcrossing the worm four times. Both CS1 and CS2 did not have any mutations after outcross (Fig. 47). In addition, RNAi of genes found by comparative genome hybridization could not establish survivors (Fig. 48). Also in NS survivors, two times outcross was performed. And WGS was performed by mixing independent six F2 progenies. In this approach, any mutation linkage point was not found (Fig. 49). Therefore,

specific mutations in functional genes are not the cause of TALT insertion by EMS.

If damage induced by EMS was not repaired, it could induce DNA damage. Thus, DNA-damaging activity of EMS might be another possibililty of inducing TALT duplication. Interestingly, telomerase-independent survivors could be established by gamma irradiation, which induces double strand breaks. Telomere of gamma irradiation-induced survivor also showed TALT insertion pattern in FISH and Southern blot (Figs. 50, 51). In addition, copy number of TALT was increased (Fig. 52). These results suggest that TALT duplication can be induced by DNA damage.

Unexpectedly, after many generations of EMS treatment, survivors still have increased mRNA expression of DNA damage response genes (Fig. 53). If normal telomeres were introduced to TALT survivors by mating, normal telomeres were changed to TALT after 10 generations without additional stimulus (Fig. 54). This features might enable the self-sustaining potential of TALT.

To find genes that regulate TALT duplication, forward and reverse genetic screening was tried. In forward genetic screening, EMS was treated in CS survivors and 2000 F2 worms were tested. After five generations, four strains have decreased TALT levels and could not be maintained (Fig. 55). However, after twenty generations, it could not found worms having reduced brood size. If mutations of TALT regulators had emerged, the brood size of survivors might be gradually reduced because CS survivors have long telomeres. However, these results suggest that mutations of TALT regulators may rapidly reduce TALT levels and brood size.

In reverse genetics screening, the first set of candidates was DNA damage and recombination-related genes. I suppressed 52 genes in CS survivor by RNAi experiments (Table 3) but I did not find any genes that could induce sterile phenotype in CS survivor. The second set of candidates was meiotic regulators. In human ALT cells, meiotic factors are adopted for homology search of short telomeres (Cho et al., 2014). I suppressed 31 genes in CS survivor by RNAi

experiments (Table 4) but these genes also were able to induce sterile phenotype in CS survivor while TALT copy numbers did not change (Fig. 56).

C-circles are a well-known ALT marker of human cancer (Henson et al., 2009). Also in *C. elegans*, *trt-1;pot-1* and *trt-1;pot-2* mutants that does not have TALT insertion show increased C-circles (Lackner et al., 2012; Shtessel et al., 2013). But C-circles levels did not increase in CS survivors (Fig. 57). Therefore, TALT survivors may not use C-circles for telomere lengthening and use a mechanism distinct from that of *trt-1;pot-1* and *trt-1;pot-2* mutants.

Discussion

1. Unique structure of TALT sequences

Based on the results, I propose a two-step model for TALT utilization in ALT. The first step involves a *cis*-duplication of a TALT donor to the subtelomere region of same chromosome to create a TALT reservoir before ALT activation, and the second involves a *trans*-duplication of telomere-adjacent TALT to all chromosomes ends via the ALT pathway. I have demonstrated that this first step event has occurred in wild strains, with TALT from a specific chromosomal 'donor' region being duplicated in *cis* to the end of the same chromosome. This telomere-adjacent copy (TALT R) acts as a reservoir, so when crisis occurs in individuals, rare survivors recruit the TALT R sequences to all chromosomal ends (*trans*-duplication), stabilizing their chromosomes.

Although I have shown for the first time that internal genomic regions can be integrated into the telomeres resulting in elongation and maintenance of telomeric ends in the absence of functional telomerase, the mechanism by which internal TALT regions are recruited and replicated to the telomeric ends by *cis*- or

trans-duplication still remains unknown. Unfortunately, I were unable to identify a single factor that suppressed the ALT phenotype in our survivors using a candidate RNAi approach.

Although expression of TALT-unique sequence is slightly increased, I do not expect that the translated protein of these sequences may have any functional significance in ALT mechanism. However, I cannot exclude the possibility that TALTs transcript can be used as templates for reverse transcription, as in Drosophila (Mason and Biessmann, 1995). Another possibility is that the unique sequence of TALT could be used as a docking site of specific proteins such as heterochromatin factors or DNA binding proteins, as in fission yeast and fly (Fanti et al., 1998; Jain et al., 2010). In these organisms, heterochromatin feature of their repetitive units is important for telomere maintenance. Coding information itself might not be important for telomere maintenance.

Flanking sequences, the other component of TALT, include canonical and variant telomere repeats. Variant telomere repeats are an emerging feature of human ALT telomere (Conomos et al., 2012; Lee et al., 2014). After whole

genome sequencing analysis of various cancer cells, only ALT positive cancers were shown to have a higher incidence of telomere variant repeats (Lee et al., 2014). In other words, in ALT cells, the canonical telomere sequence is replaced by variant telomere sequences in an interspersed pattern. The insertion of variant telomere is thought to alter telomere-binding proteins and recruits nuclear orphan receptors (COUP-TF2 and TR4) and nucleosome remodeling deacetylase (NuRD complex) (Conomos et al., 2014; O'Sullivan and Almouzni, 2014). I anticipate that variant repeats of TALTs may also recruit other binding proteins, as in human ALT cells.

2. DNA damage response: a possible mechanism of *trans*-duplication.

The DNA damage response (DDR) should be suppressed within telomere to avoid chromosome fusion and to maintain integrity. In normal mammalian telomeres, shelterin complexes perform this protective role. However, in ALT

tumor cells, the density of shelterin is decreased, and variant repeat binding proteins cannot fully suppress DDR (Cesare et al., 2009). These altered telomere proteins reportedly suppress non-homologous end joining, but not homologous recombination, thereby allowing ALT. The imperfectly protected telomere state described above is called the ‘intermediate’ state of telomeres.

It is conceivable that TALTs in worms may recruit binding proteins that differ from canonical telomere-binding proteins, as in ALT tumor cells. Two piece of evidences show that TALTs may induce a condition similar to the ‘intermediate’ state of telomere. First, the mRNA expression of DDR genes is specifically up-regulated in TALT survivors. Second, TALTs integrated into telomeres can be self-sustained without additional stimulus, even when canonical telomere sequence are provided by mating ALT survivors with telomerase positive worms. Therefore, TALTs at telomeres are likely to be in an intermediate state that could induce transposition or recombination.

3. Two types of ALT mechanisms in *C. elegans.*

Unlike unstably-maintained chunking survivors that utilized only telomere repeats, TALT survivors can be stably maintained. Chunking survivors have a large amount of C-circles, which are extrachromosomal, single-stranded C-rich telomeric circular DNA molecules (Cheng et al., 2012; Shtessel et al., 2013). It might be a template of ALT recombination, thus indicating a recombination-dependent ALT mechanism.

However, in my study, TALT survivors use complex tandem repeats instead of simple telomere repeats to cope with telomere loss. Therefore, there are at least two distinct telomere lengthening mechanisms in *C. elegans*: telomere-telomere recombination by C-circle and formation of complex tandem repeats using non-telomeric sequence. In nature, many organisms use non-telomeric tandem repeats to preserve telomere. TALT insertion is reminiscent of the evolutionarily conserved mechanism that is repeatedly accepted to overcome telomerase loss.

4. *Cis*-duplication of TALT in subtelomeres evolution.

The telomerer-adjacent regions of N2 and CB4856 have been independently duplicated with their own TALT during evolution. In other organisms, the telomere-adjacent regions are called subtelomeres (Riethman et al., 2005). Subtelomeres are highly polymorphic regions because they have high rates of mutation and recombination (Linardopoulou et al., 2005). Although the definition of subtelomeres is not intuitively clear, the subtelomeres of many organism have certain feature in common. Subtelomeres are composed of various repeat elements, including variant telomere repeats, but the extent of repeat duplication and divergence varies greatly within and between species (Mefford and Trask, 2002). Because of their highly variable and repetitive features, the complete assembly of the subtelomeric regions has not been achieved thoroughly. Therefore, the structure and function of subtelomers have remained mainly unanswered.

Subtelomeric regions can be used as templates for telomere lengthening under telomerase-deficient conditions. For example, Y' element of budding yeast that can be utilized for type I survivors is located in subtelomere (Lundblad and Blackburn, 1993). The fission yeast *Schizosaccharomyces pombe* uses the subtelomere as a template for ALT activation (Nakamura et al., 1998). Additionally, in human ALT cells, telomere variant repeats are enriched in subtelomeres (Varley et al., 2002). TALT amplification clearly shows that subtelomeric regions can be amplified for ALT activation, perhaps representing first description of this phenomenon in a multicellular organism. I speculate that TALTs may be found in other organisms, including humans, if read lengths and the cost of whole genome sequencing are improved so as to obtain a completely assembled map of subtelomere.

First telomere study of *C. elegans* was started in N2 background. Following this research, researchers described *C. elegans* telomere sequence but argued that *C. elegans* does not have subtelomeres. But in my research, I found multiple copies of TALT1 in telomere-adjacent region in the right arm of

chromosome V in CB4856. It is originated from an internal region of the same chromosome by *cis*-duplication. This duplication event might be similar to the subtelomere formation process of other organisms.

Subtelomeric regions in humans rapidly obtained unique sequences during primate evolution (Baird and Royle, 1997). Likewise, I observed in *C. elegans* that a sequence near the proximal telomere, TALT, appears to be rapidly changing, as it has been independently duplicated in multiple wild isolates. Therefore, I propose that a rapidly evolving proximal telomere region retains highly replicative potential, so that it can be preferentially selected as a template for ALT at telomere crisis. Bioinformatics analysis identified more TALT-like sequences in the *C. elegans* reference genome. It would be of interest to examine the wild isolates that lack TALT1 and TALT2 in order to assess whether they utilize other TALT-like sequences for ALT (Fig. 58 and Table 5). In addition, phylogeny trees using proximal telomere region show TALT having strains can be tied by bioinformatics analysis (Figs. 59, 60 and Table 6).

During evolution of linear chromosome, one possible explanation is that symbiont's mobile group II introns having RT sequence might invade circular genome and cut the genome to a linearized form (De Lange, 2015). After cutting, these mobile elements might stabilize ends by repetitive invasion. The *cis*-duplication of TALT might be another possible mechanism for cutting of circular DNA by moving internal elements. After *cis*-duplication, it can stabilize emerging ends by *trans*-duplication. Interestingly, in phylogenic tree of *C. elegans* wild isolates, *cis*-duplication of TALT2 has occurred independently. In nature, some kinds of environmental condition might be selection pressure for TALT cis-*duplication*.

5. Remaining questions.

In evolutionary process, the simultaneous appearance of linear chromosome and telomerase would not happen by accident (Fajkus et al., 2005). Before the advent of telomerase, other mechanisms like recombination and

retrotransposition, which were already present in prokaryote, could had been established (Nosek et al., 2006). These mechanisms could be re-activated after telomerase defect. Telomerase loss can make organism revert back to previous state and recall ancient mechanism like *Diptera*.

I found that internal genomic regions, TALTs, can be duplicated to telomeres and stabilize the genome after telomerase loss in multicellular organisms. TALT duplication might represent the re-activation of an ancient mechanism that existed before telomerase evolved.

Before activation of TALT duplication, chromosomes could be fused by telomere defects. The altered karyotype can still be stably maintained after TALT insertion. Thus, TALT survivors might suggest one partial case of chromosome evolution.

TALTs contain unique genomic region flanked by telomere-like sequences. TALT-like DNA structures also exist in its genome of SV40, which is used for cancer transformation (Fasching et al., 2005; Marciniak et al., 2005). These sequences can also be duplicated to telomeres during tumor transformation.

Hence, TALTs may represent an evolutionarily conserved structure for telomere maintenance. However, I do not know which parts are important for TALT movement. Therefore, I am trying to identify the necessary parts of TALT for movement, such as unique sequence of TALT or the telomere-like repeats, via the clustered regularly interspaced short palindromic repeat (CRISPR) system.

The *trans*-duplication of TALTs can overcome telomerase defect. Currently, the mechanism of *trans*-duplication remains unknown. I hypothesize that TALTs may have other binding proteins in addition to telomere binding proteins. These proteins may initiate some type of signal that recruits the duplication machinery. Proteomic approaches will enable the identification of TALT-binding proteins. In addition, to identify the duplication machinery, I conducted candidate RNAi experiments with recombination and DNA damage response factors. However, I did not identify the TALT duplication machinery in that screen, thus suggesting that an unknown machinery may regulate the *trans*-duplication. Therefore, I am now trying to identify possible clue with a forward genetic screening.

The *cis*-duplication of TALT can occur without telomerase loss. *cis*-duplicated TALT can be used for templates of *trans*-duplication in dangerous situations. Interestingly, in phylogenic tree of wild *C. elegans* isolates, *cis*-duplication of TALT has independently occurred many times. In nature, some kinds of environmental condition would make selection pressure for TALT cis-*duplication.* Therefore, *cis*-duplicated TALT is ‘scar’ of natural selection and a reservoir for *trans*-duplication. This duplication event might be similar to the subtelomere formation process of other organisms. Consistently, in humans, duplicated units in the subtelomeres may be used for ALT cancer formation. Therefore, understanding mechanisms underlying TALTs will provide valuable information for ALT cancer therapy.

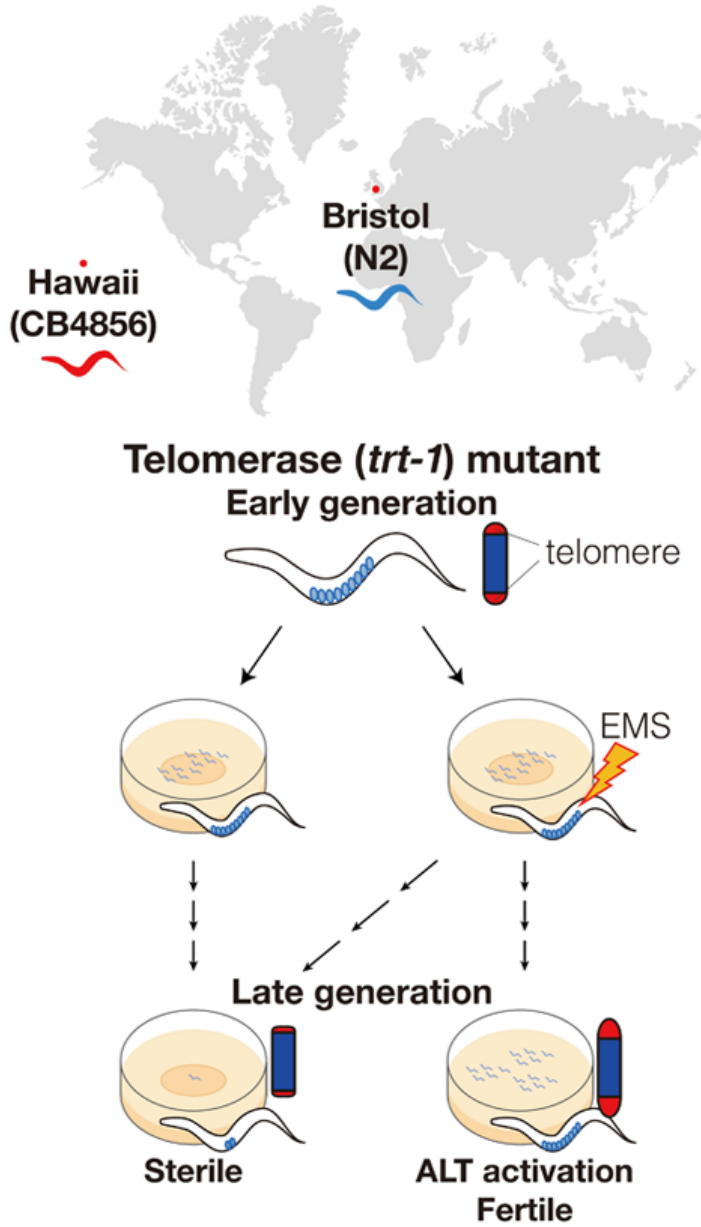


Figure 1. Isolation of stable ALT survivors in *C. elegans*. A schematic diagram showing the experimental procedures to isolate ALT survivors in two wild. In collaboration with B. Seo and D. S. Lim

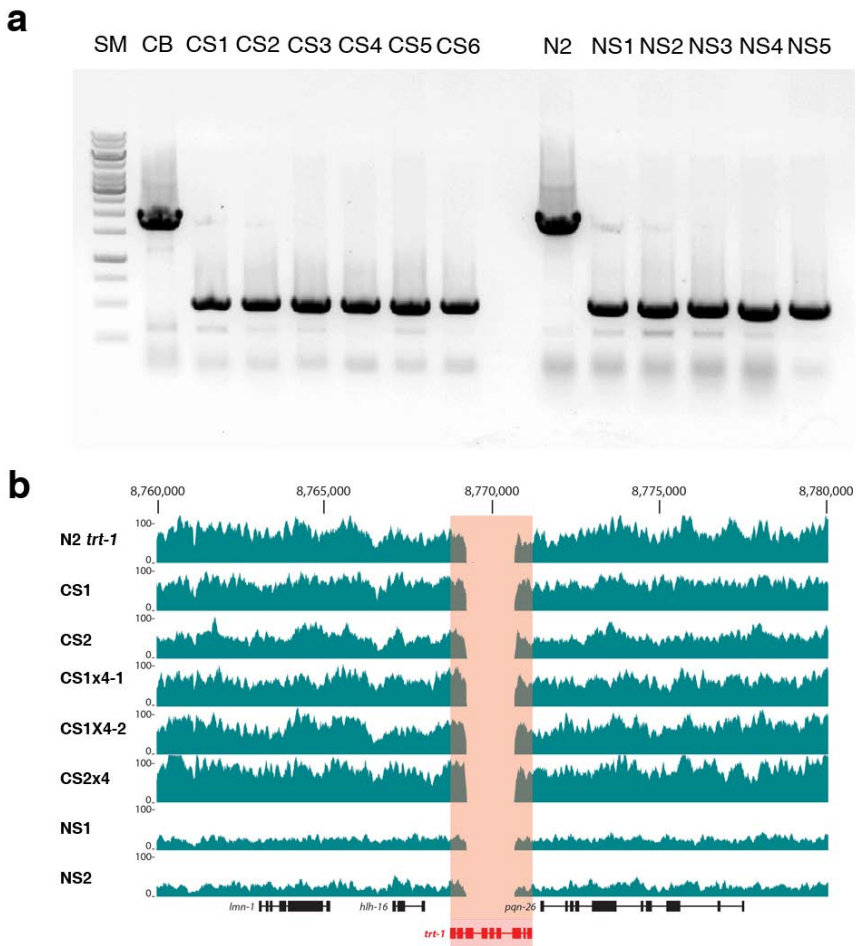


Figure 2. ALT survivors have homozygote telomerase mutation. (a) PCR amplification of the *trt-1* deletion allele. The *trt-1(ok410)* deletion mutation results in PCR amplicon shorter than wild-type control. CB, CB4856. **(b)** Coverage plot of whole genome sequencing reads around the *trt-1* locus confirms the deletion

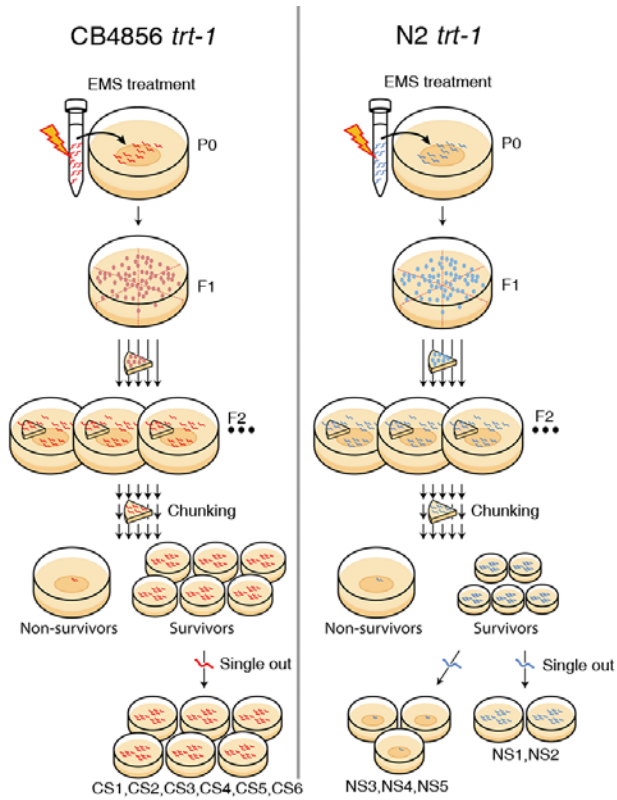


Figure 3. A schematic diagram of the isolation protocol for ALT survivors.

For CS survivors, worms were separated to 80 plates after EMS treatment. After 8 generations, 6 resulting survivors (CS1-CS6) were maintained by transfer of small number of worms each generation. For N2 survivors, worms were separated to 200 plates after EMS treatment. 5 survivors (NS1 – NS5) were maintained by large chunking, two of which (NS1 and NS2) were subsequently maintained by transfer of small number of worms each generation. In collaboration with B. Seo and D. S. Lim

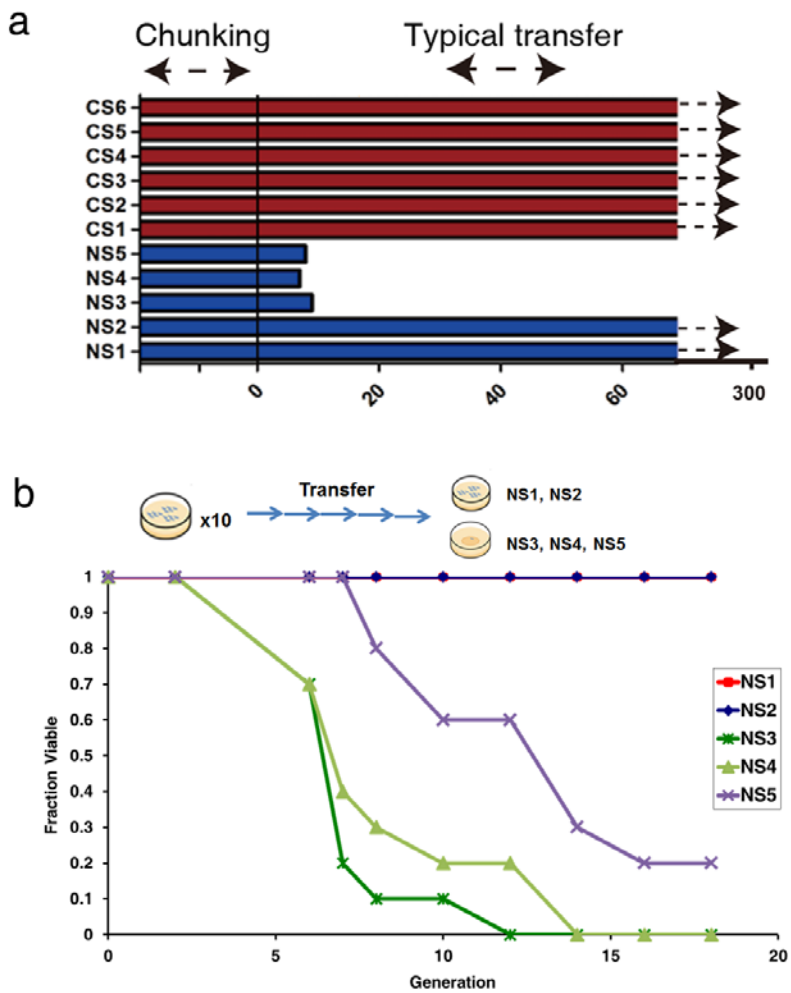


Figure 4. Survival graph of ALT survivors. (a) 6 CS survivors and 2 NS survivors could be maintained for at least 300 generations by transferring 10-15 L1 larvae. **(b)** Survival graph of NS survivors**.** 10 plates of NS3, NS4 and NS5 reached sterility within 20 generation by transferring 10-15 larvae. In collaboration with B. Seo and S. Sung.

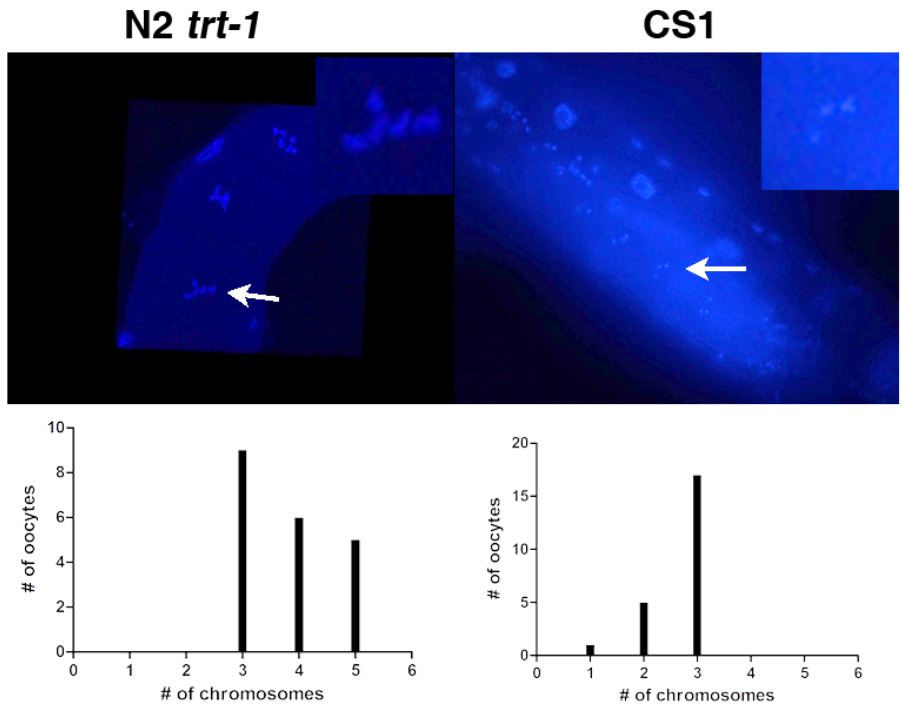


Figure 5. The chromosome number of CS survivor was decreased. Histogram of chromosome numbers of N2 *trt-1* and CS1 in diakinesis of the germline. Top right corner shows the magnified images of nuclei indicated by arrows. Total n= 20, 21 respectively. In collaboration with S. Sung.

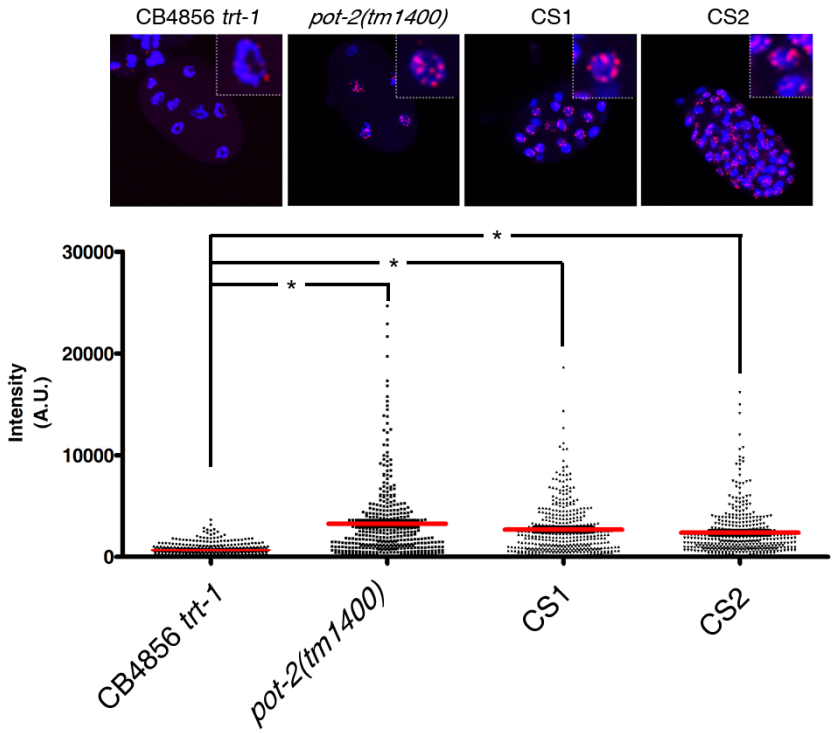


Figure 6. Telomere lengths of CS survivors were increased in FISH. Quantitative FISH analysis for telomere length of the survivors CB4856 *trt-1*, *pot-2(tm1400)*, CS1 and CS2. Telomere was detected by Cy-3-TTAGGC*3 PNA probe (red) in the embryo. DNA was counterstained with DAPI (4',6-diamidino-2-phenylindole) (blue). The upper panel shows representative images for each strain. T-test was used for statistical analysis for quantification (*P-value < 0.0001, n = 388/each). Mean value is represented with red bar. In collaboration with B. Seo.

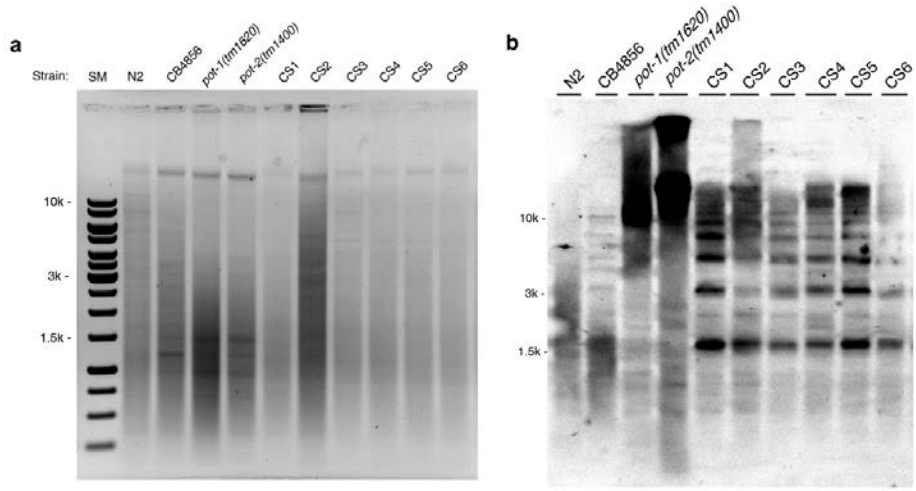


Figure 7. Telomere lengths of CS survivors were increased in TRF. (a) Ethidium bromide stained gel image and uncropped image **(b)** TRF analysis of N2, CB4856, *pot-1(tm1620), pot-2(tm1400)* and all the CS survivors digested with a combination of 6-cutter restriction endonucleases (*Nhe*I, *Dra*I, *Apa*I, *Nde*I, *Xho*I, *Nco*I and *Sac*I) and probed with DIG-TTAGGC*4. *pot-1(tm1620)* and *pot-2(tm1400)* were used as positive controls as they have long telomeres.

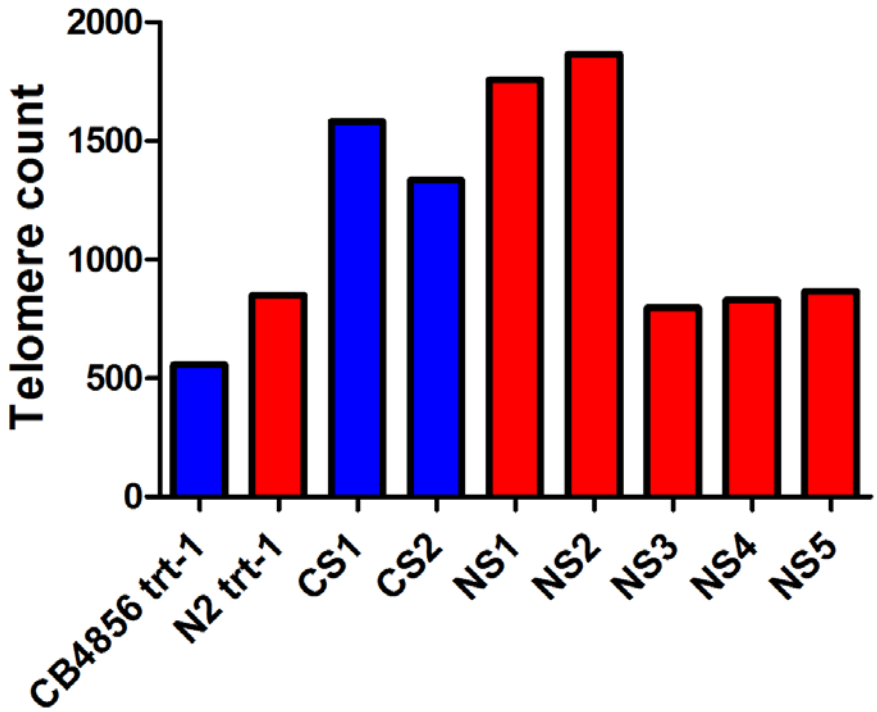


Figure 8. Telomere lengths of ALT survivors were increased. Normalized count of reads containing at least 6 telomere repeats in CB4856 trt-1, CS1, CS2 and NS1-NS5 worms. In collaboration with M. Hills.

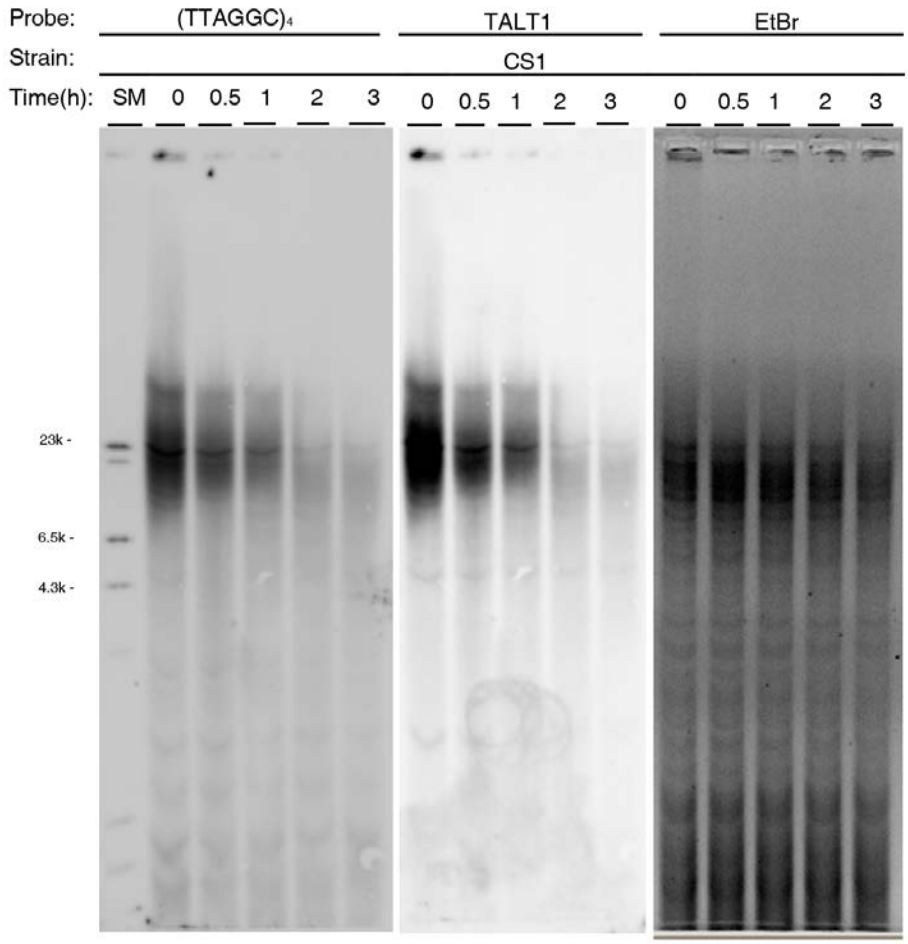


Figure 9. Telomere length was increased in CS survivor at the chromosome ends. BAL 31 exonuclease assay of CS1. Genomic DNA was treated by BAL 31 exonuclease prior to digestion with TALT non-cutting restriction enzyme mix. Digested DNA was then analyzed by Southern blot. In collaboration with S. Sung.

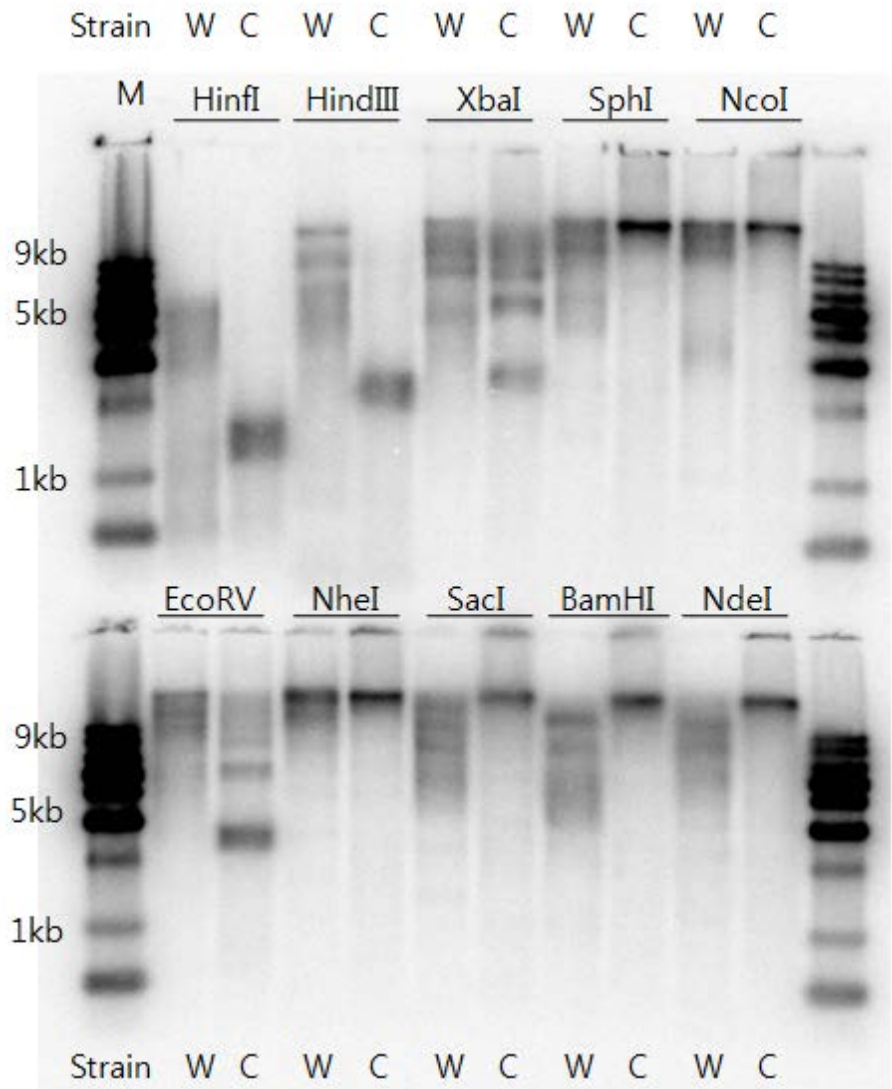


Figure 10. CS survivor show distinct TRF patterns. TRF analysis using various restriction enzymes that does not cut canonical telomere sequence. The blot was hybridized with DIG-TTAGGC*4. W, wild type; C, CS1.

DpnI	HaeIII	HpaII	MseI	NlaIII	RsaI
AatII	DraI	KpnI	HinfI	NarI	XbaI
ApaI	EcoRI	Sac	NaeI	SacII	XhoI

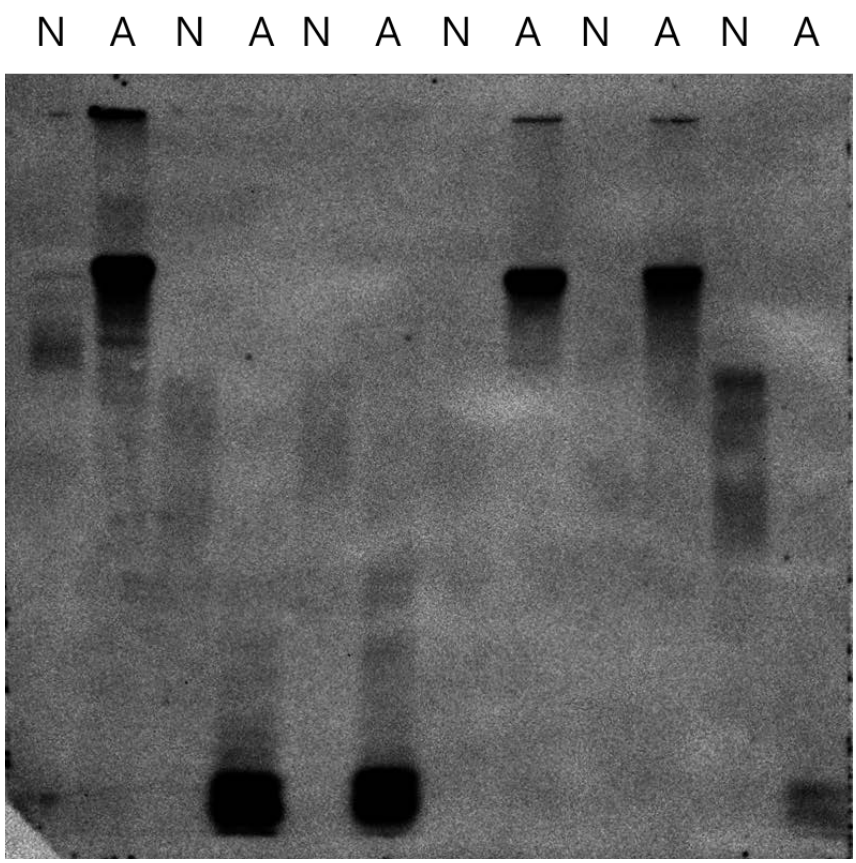


Figure 11. NS survivor show distinct TRF patterns. TRF analysis using various restriction enzymes that does not cut canonical telomere sequence. Upper panel show restriction enzyme list. The blot was hybridized with DIG-TTAGGC*4. N, N2; A, NS1 survivor.

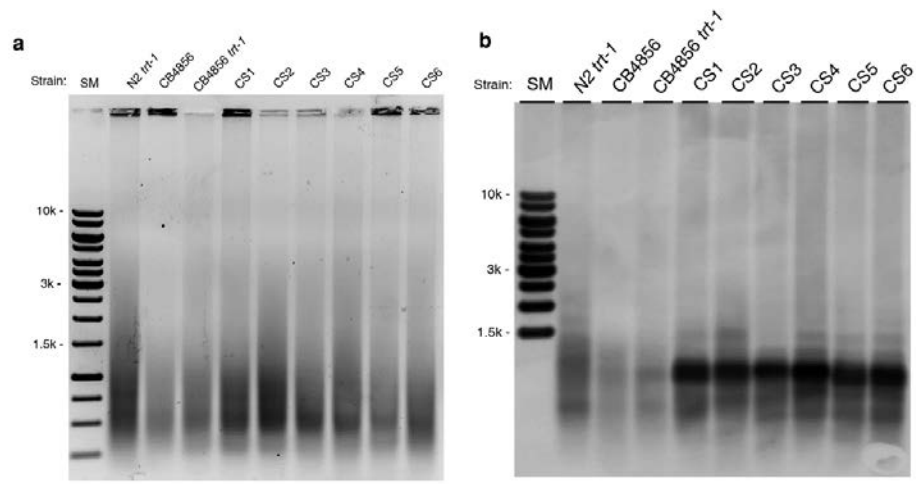


Figure 12. CS survivors have distinct telomere sequences. (a) Ethidium bromide stained gel image and uncropped image (b) TRF analysis of N2 *trt-1*, CB4856, CB4856 *trt-1* and all the CS survivors probed with DIG-TTAGGC*4 after *Hin*fI digestion.

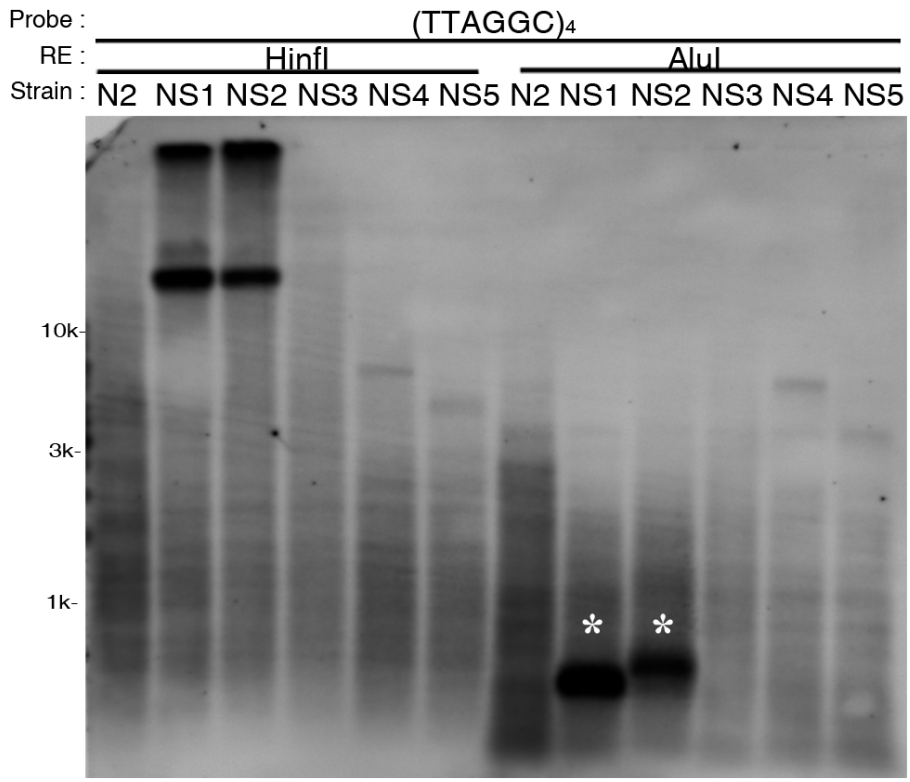


Figure 13. Only the NS1 and NS2 show discrete banding patterns. TRF analysis for telomere of N2 survivors. Genomic DNA of N2 *trt-1* and NS survivors (NS1 to NS5) was digested with *Hin*fI or *Alu*I. The blot was hybridized with (TTAGGC)*4 repeat. Asterisks indicate cleaved telomeres.

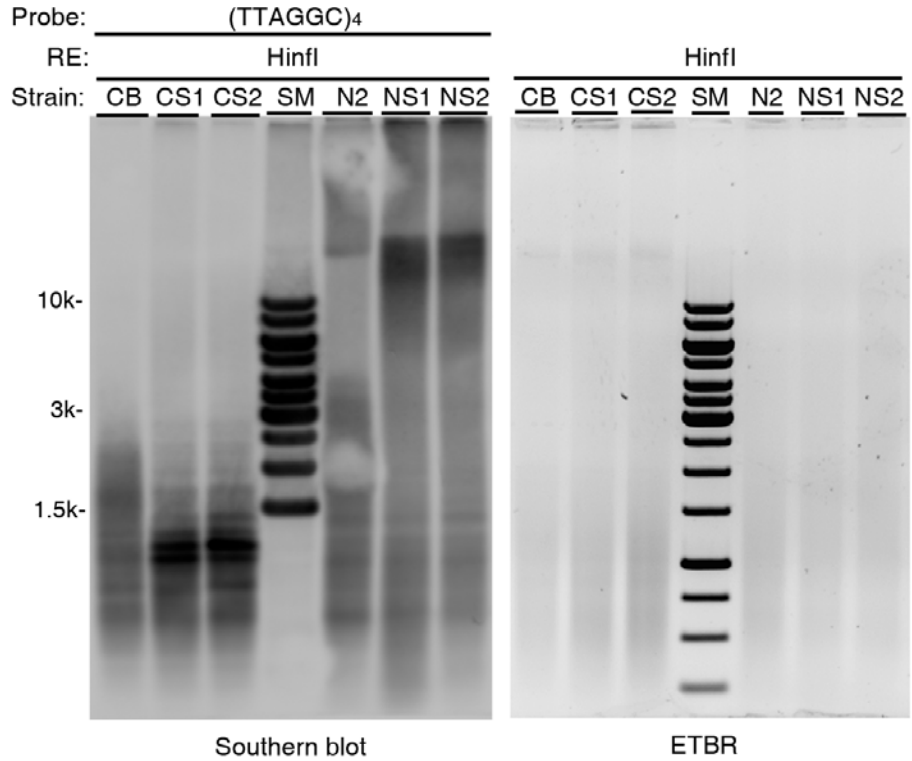


Figure 14. ALT survivors utilize different TALT loci in a strain-dependent manner. CS and NS survivor show different TRF pattern. The blot was hybridized with DIG-TTAGGC*4 after *Hin*fI.

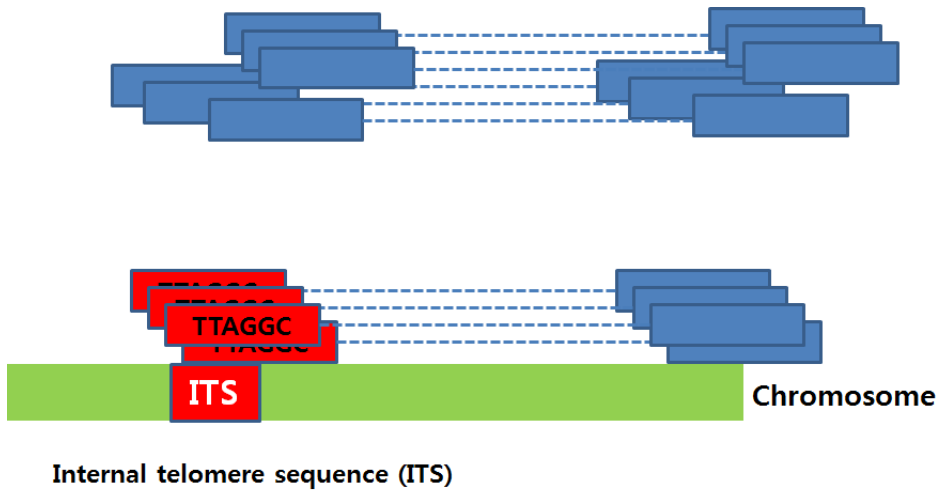


Figure 15. A schematic diagram of paired-end reads analysis. Collecting telomere-containing reads and aligned to reference N2 genome.

In collaboration with M. Hills.

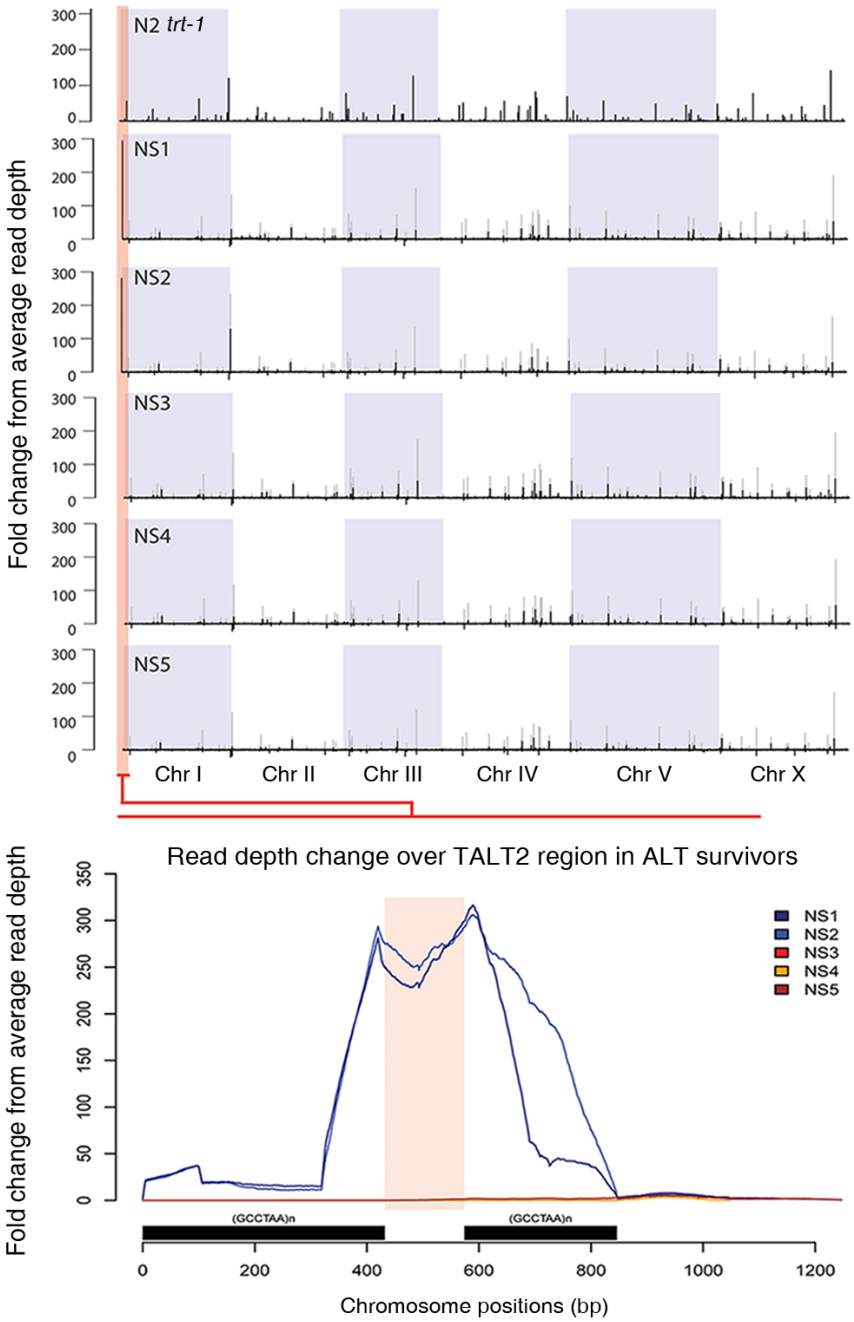
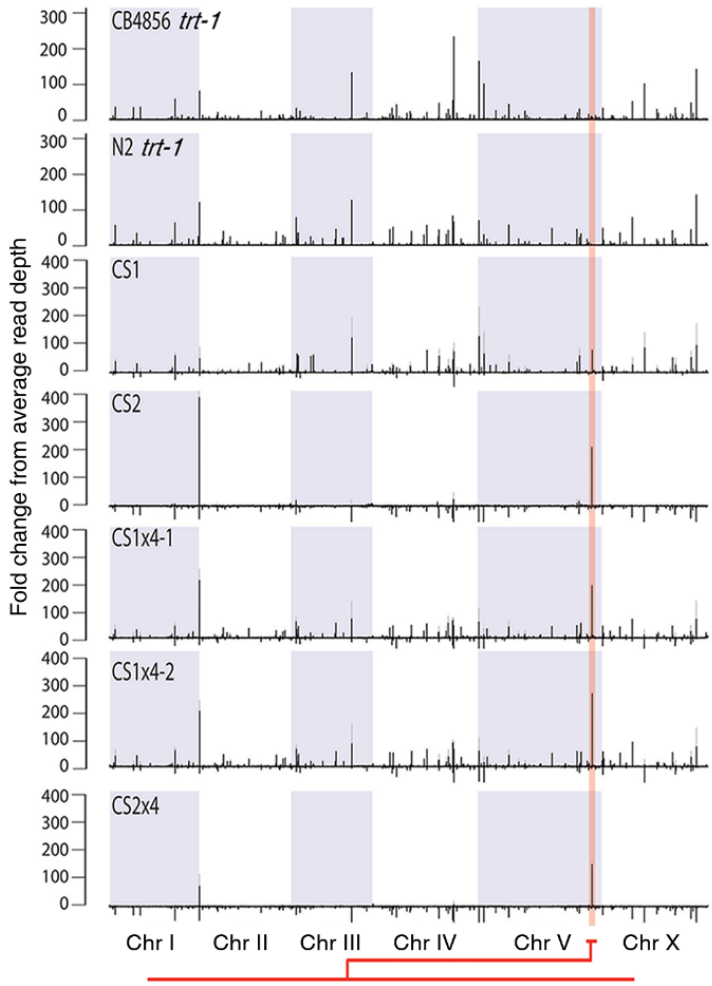

N2 trt-1
NS1
NS2
NS3
NS4
NS5
300
200
100
0
Fold change from average read depth
Chr I
Chr II
Chr III
Chr IV
Chr V
Chr X
Read depth change over TALT2 region in ALT survivors
NS1
NS2
NS3
NS4
NS5
350
300
250
200
150
100
50
0
(GCCTAA)n
(GCCTAA)n
0
200
400
600
800
1000
1200
Chromosome positions (bp)

Figure 16. Amplified region in NS survivors. Fold-change from average read depth plot (black bar) of N2 *trt-1* and N2 survivors. To normalize strain specific change, fold-change from average depth from parental N2 *trt-1* was subtracted from total fold-change (grey bar) of ALT survivors. The bottom panel is an enlargement of TALT on chromosome I. Internal telomere repeats are indicated as black bars. In collaboration with M. Hills.



Read depth change over TALT1 region in ALT survivors

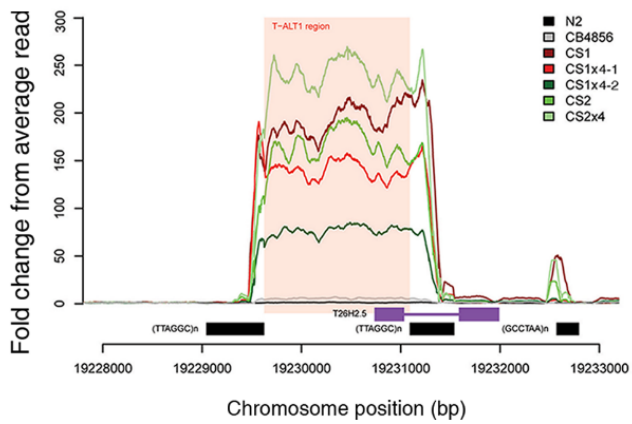
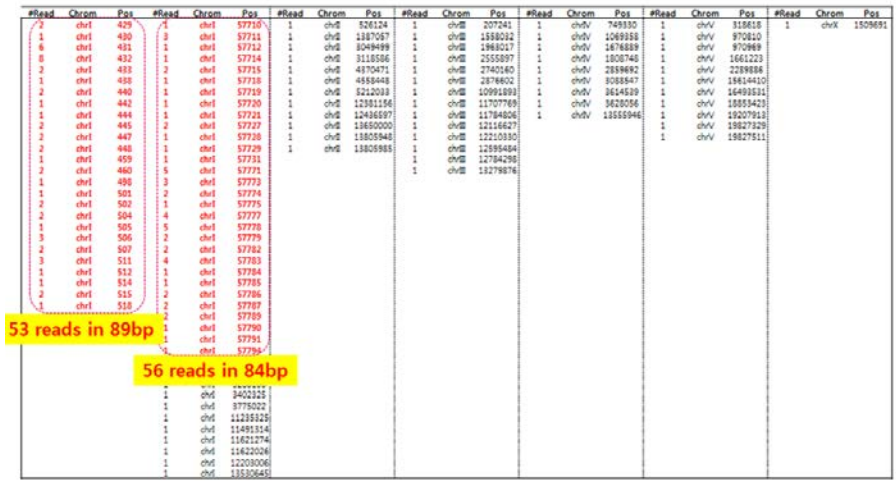


Figure 17. Amplified region in CS survivors. Fold-change from average read depth plot (black bar) of CB4856 *trt-1*, N2 *trt-1*, CS1, CS2 and outcrossed CS survivors with N2 (CS1x4-1, CS1x4-2 and CS2x4). To normalize strain specific change, fold-change from average depth of parental CB4856 *trt-1* was subtracted from total fold-change (grey bar) of ALT survivors. While positive value indicates over-representation of the sequence, negative value indicates under-representation of the sequence compared to control. The bottom panel is an enlargement of TALT on chromosome V (red shade). Internal telomere repeats are indicated by black bars. Gene structure flanking TALT1 locus is denoted by purple bar. In collaboration with M. Hills.

NS1



trt-1(ok410);pot-1(tm1620)

#Read	Chrom	Pos	#Read	Chrom	Pos	#Read	Chrom	Pos	#Read	Chrom	Pos	#Read	Chrom	Pos
1	chrI	57726	1	chrII	337010	1	chrIII	209334	1	chrIV	1141910	1	chrV	232458
1	chrI	256125	1	chrII	526014	1	chrIII	1330717	1	chrIV	1141911	1	chrV	829769
1	chrI	727168	1	chrII	618679	1	chrIII	1753511	1	chrIV	1808759	2	chrV	970961
1	chrI	775517	1	chrII	645121	1	chrIII	2740271	1	chrIV	3166496	1	chrV	999691
1	chrI	802203	1	chrII	3043276	1	chrIII	4950902	1	chrIV	3763011	1	chrV	1305824
1	chrI	968681	1	chrII	3135485	1	chrIII	4987347				1	chrV	1317167
1	chrI	974723	1	chrII	3248401	1	chrIII	10784591				1	chrV	2682025
1	chrI	11465946	1	chrII	3957865	1	chrIII	10784620				1	chrV	3233722
2	chrI	11621274	1	chrII	4106499	1	chrIII	11207739				1	chrV	3641389
1	chrI	11622007	1	chrII	5212051	1	chrIII	12209341				1	chrV	4026555
1	chrI	11622015	1	chrII	11956957	1	chrIII	12801591				1	chrV	18797541
1	chrI	11622020	1	chrII	13805951	1	chrIII	13280009						
1	chrI	13927915	1	chrII	14122821									

Figure 18. Long telomere mutant did not have the amplified region in chromosome I. Upper panel show amplified region in NS survivor. Left of chromosome I amplified in NS survivor. Bottom panel show *trt-1(ok410);pot-1(tm1620)* didn't have amplified region.

1	2	3	4	5	6
DpnI	HaeIII	HpaII	MseI	NlaIII	RsaI
AatII	DraI	KpnI	HinfI	NarI	XbaI
ApaI	EcoRI	SacI	NaeI	SacII	XhoI

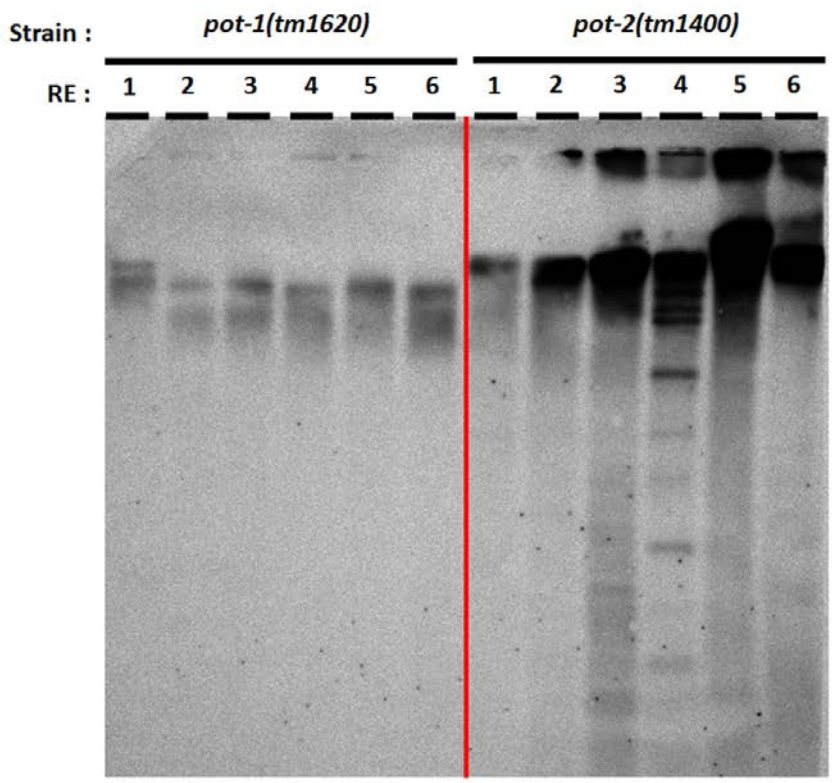


Figure 19. Long telomere mutant did not have the discrete band patterns. Both *pot-1(tm1620)* and *pot-2(tm1400)* did not show TALT insertion restriction patterns. This blot was hybridized with DIG-TTAGGC*4.

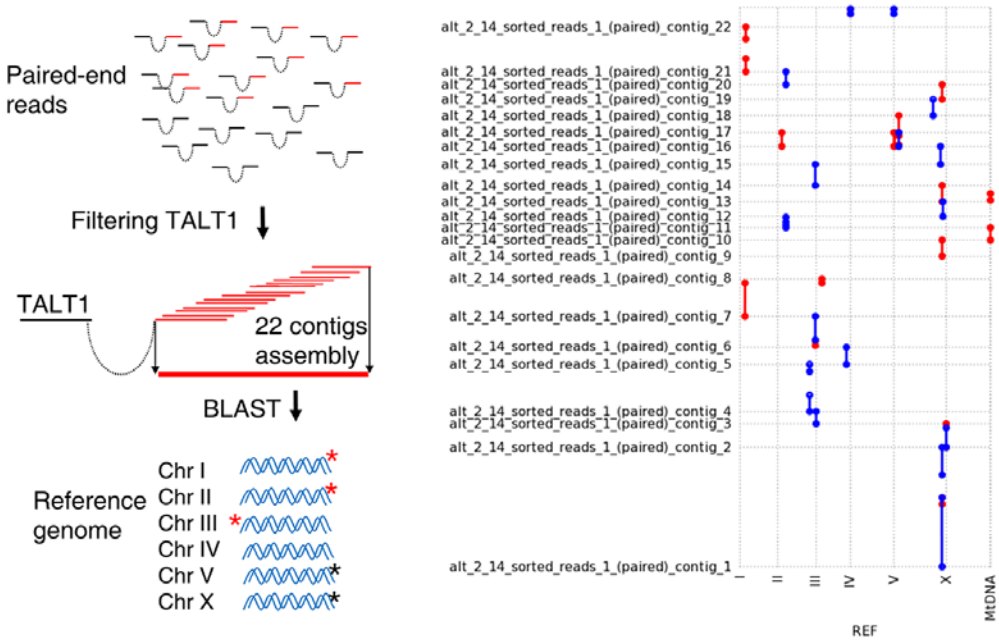


Figure 20. TALT1 was replicated in the telomere of CS survivor. Diagram illustrating how the contigs (red bar) were reconstructed from reads in which one pair aligned to TALT1. Asterisks indicate predicted breakpoint by BLAST (black) or confirmed breakpoints (red). 22 contigs were made in CS1x4. In collaboration with H. Oh.

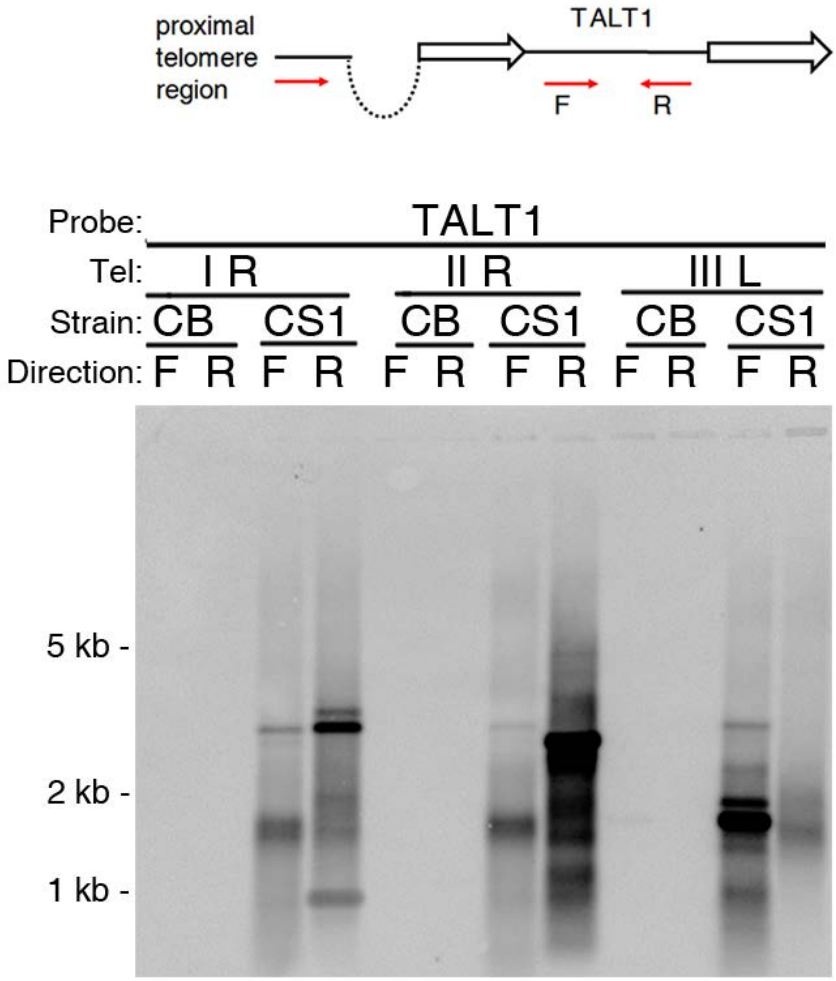


Figure 21. TALT1 was replicated in the telomere of CS survivor. Confirmation of *trans*-duplication by breakpoint PCR followed by Southern blot with TALT1 probe. Breakpoint PCR primers (red arrows) were designed to amplify the junction of TALT1 and each of proximal telomere regions. F, forward primer; R, reverse primer; open arrow, telomere repeat.

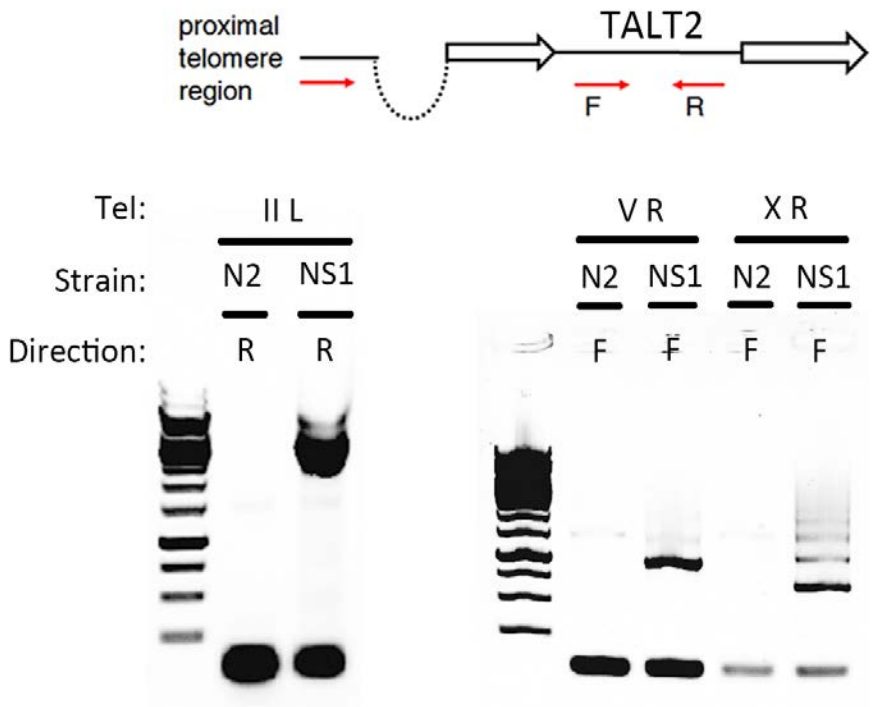


Figure 22. TALT2 was replicated in the telomere of NS survivor. Confirmation of *trans*-duplication by breakpoint PCR. Breakpoint PCR primers (red arrows) were designed to amplify the junction of TALT2 and each of proximal telomere regions. F, forward primer; R, reverse primer; open arrow, telomere repeat. In collaboration with E. Kim.

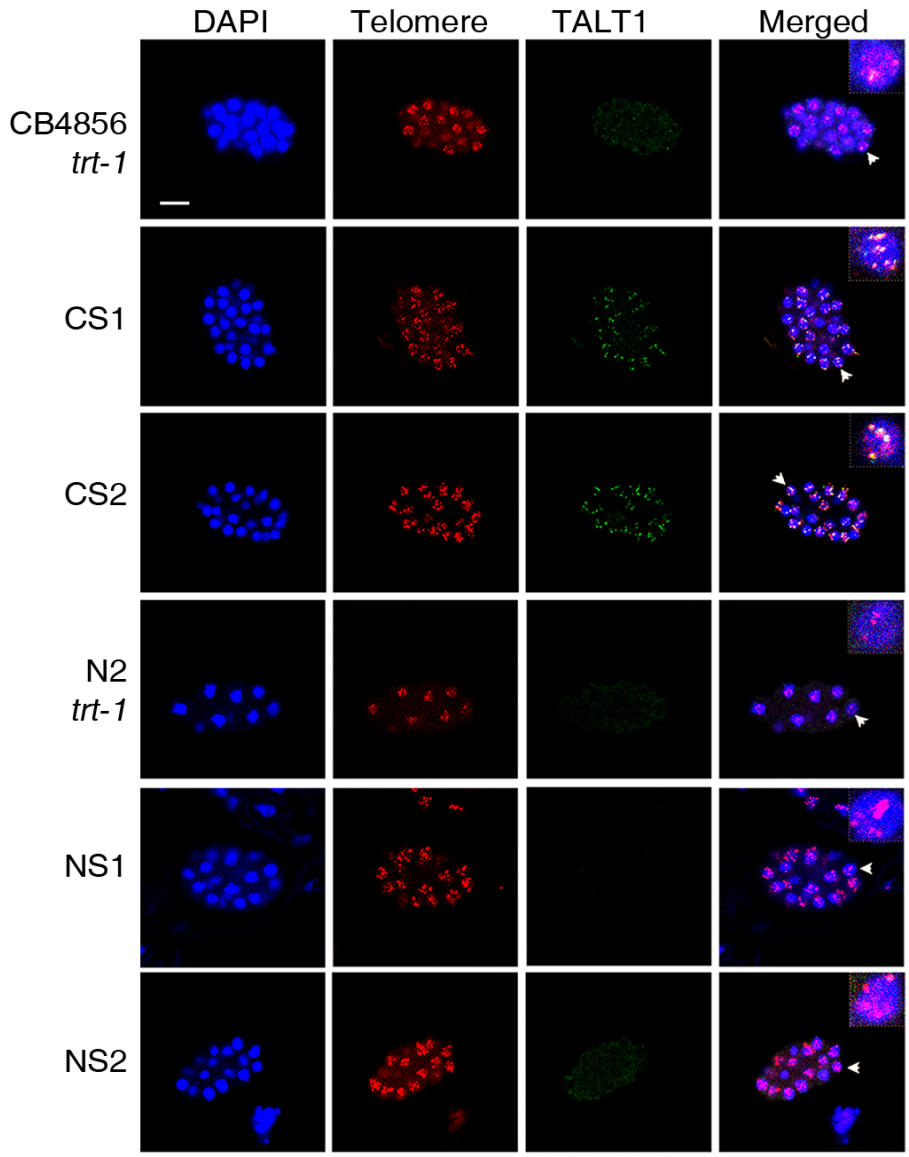


Figure 23. TALT1 is co-localized with telomere only in CS survivors. FISH using the TALT1 probe and telomere probe in embryonic stage. Scale bar, 10 μm; blue, DAPI; green, TALT1; red, telomere. In collaboration with B. Seo.

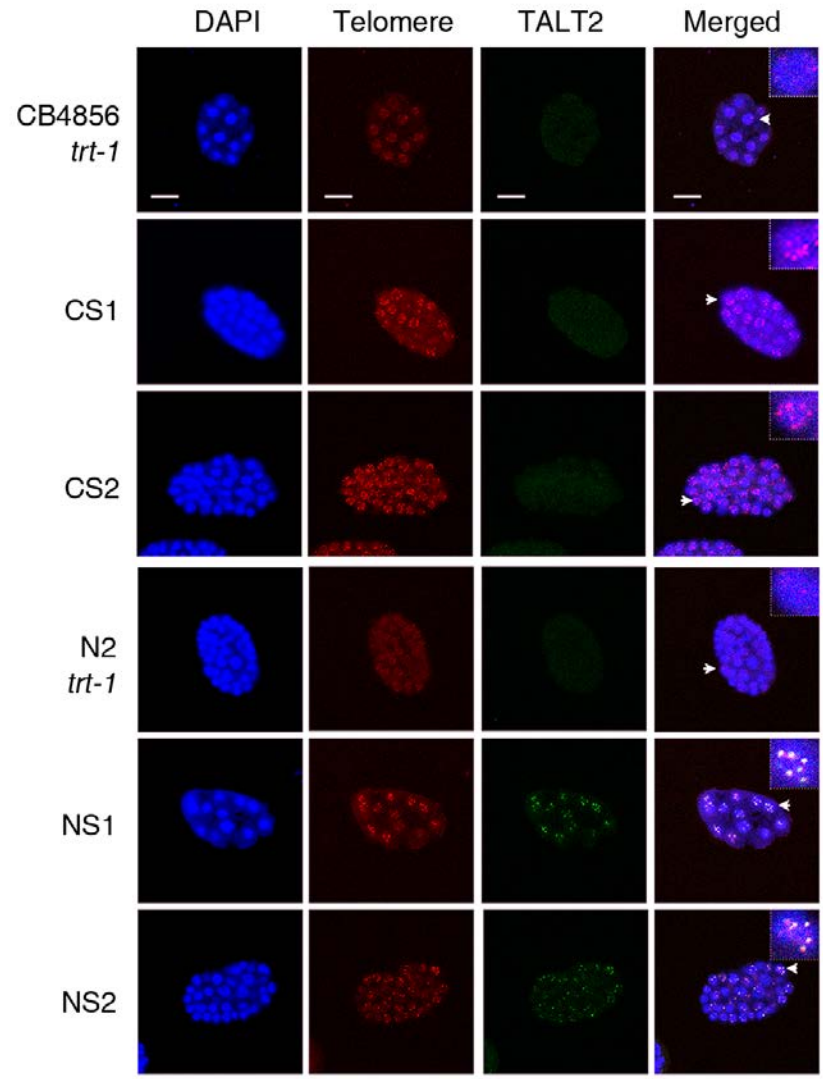


Figure 24. TALT2 is co-localized with telomere only in NS survivors. FISH using the TALT and telomere probe in embryo of CB4856 *trt-1*, N2 *trt-1* and all the survivors. Scale bar, 10 μm; blue, DAPI; green, TALT; red, telomere. In collaboration with B. Seo.

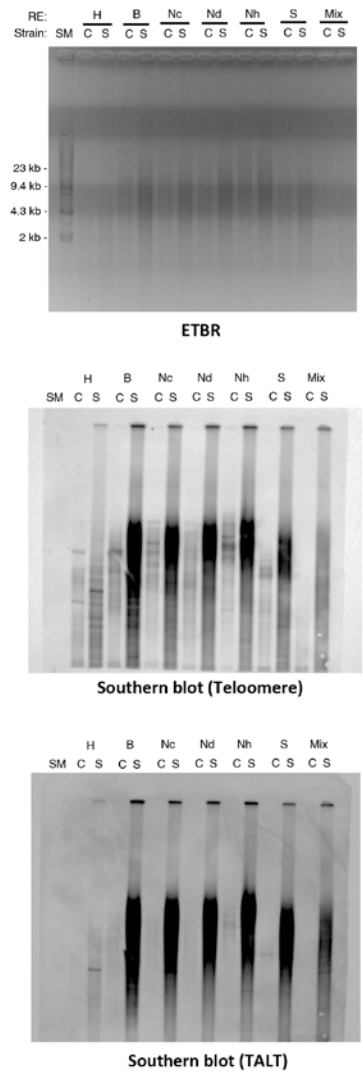


Figure 25. TALT1 have established long telomere of CS survivor. TRF analysis using a TALT1 probe and a telomere probe after TALT non-cutting enzyme digestion. H:HindIII, B:BamHI, Nc:NcoI, Nd:NdeI, Nh:NheI, S:SacI, Mix:mixure of HindIII, BamHI, NcoI, NdeI, NheI and SacI.

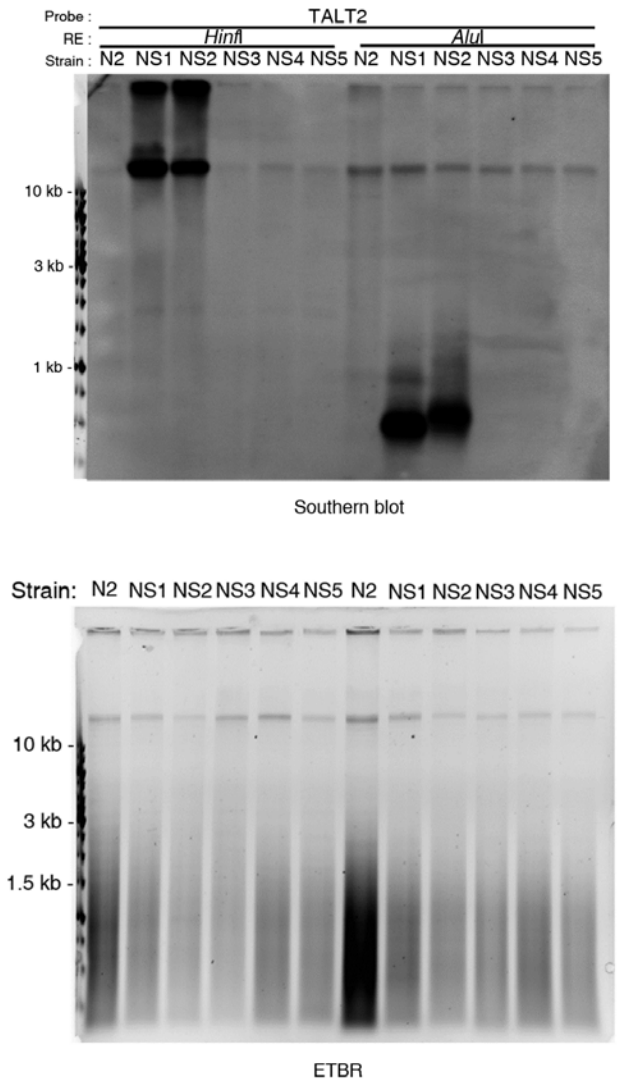


Figure 26. TALT2 have established long telomere of NS survivor. TRF analysis using a TALT2 probe after *Hinf*I and *Alu*I digestion. *Hinf*I is TALT non-cutting restriction enzyme and *Alu*I is TALT cutting restriction enzyme. The bottom panel is ethidium bromide stained gel image and uncropped image.

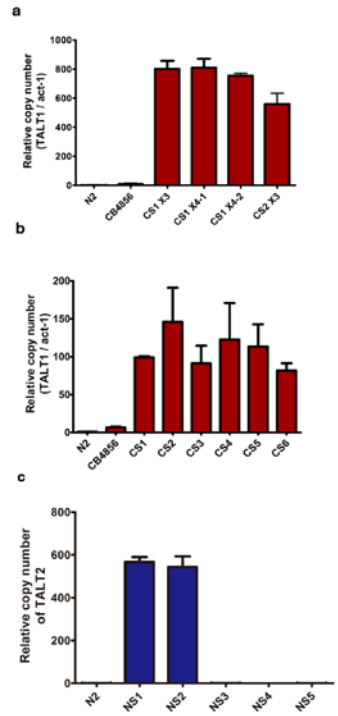


Figure 27. Telomere-independent survivors have higher TALT copy number.

(a) Copy number measurement of TALT1 by quantitative PCR (qPCR). After outcrosses to N2, TALT1 copy number increased in outcrossed CS survivors. Copy number was normalized to the *act-1* gene. n = 3. Bar is ± S.D. **(b)** Copy number measurement of TALT1 by qPCR in all CS survivors. n = 3. Bar is ± S.D. **(c)** Copy number measurement by qPCR using a TALT2-specific primer. Both stable NS survivors (NS1, NS2) have increased level of TALT2. Copy number was normalized to the *act-1* gene. n = 3. Bar is ± S.D.

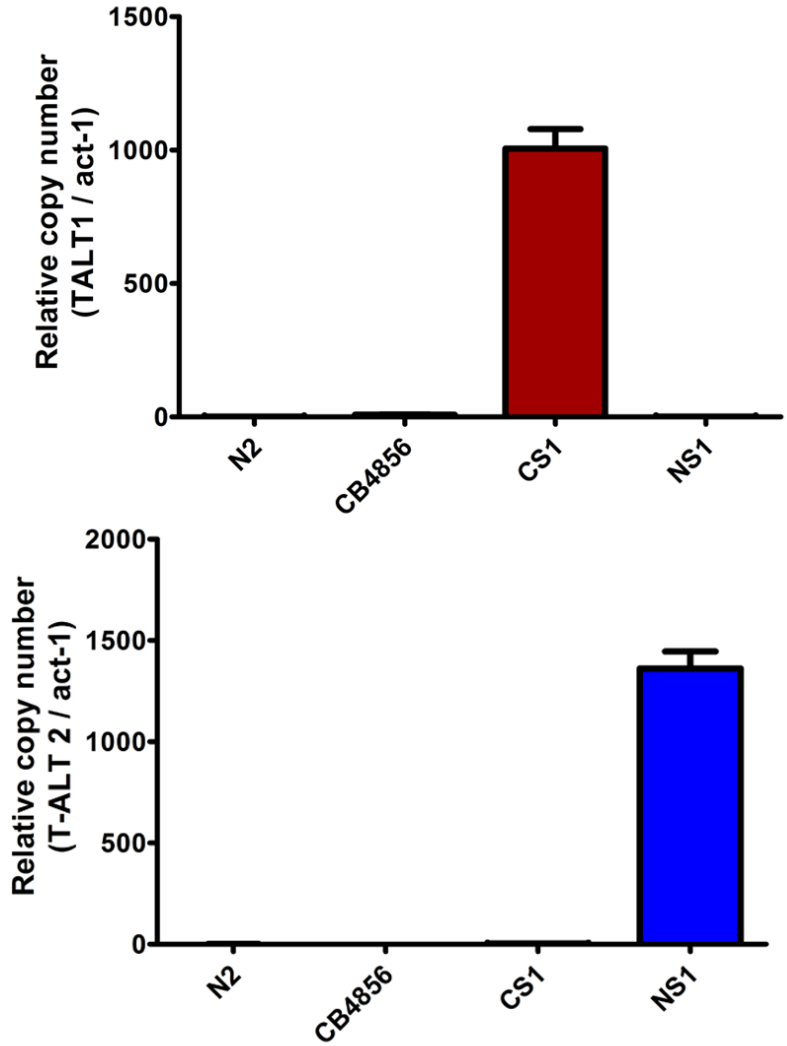
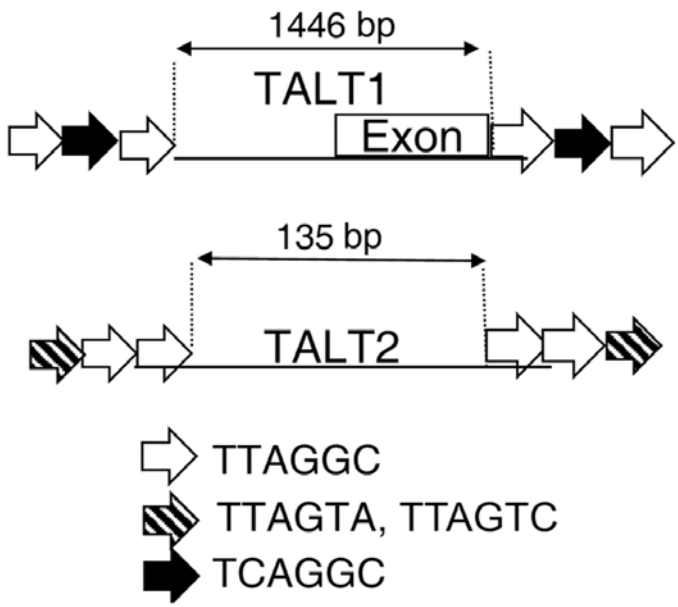


Figure 28. ALT survivors utilize different TALT loci in a strain-dependent manner. Copy number of TALT1 specifically increased in CS1, while copy number of TALT2 specifically increased in NS1, assessed by qPCR and normalized with *act-1*. *n* = 3. Bar is ± S.D.



	TALT1	TALT2
Background	CB4856	N2
Origin	Chromosome V	Chromosome I
Telomere variant	TCAGGC	TTAGTC, TTAGTA
Unique sequence length	1446	135

Figure 29. The feature of TALTs. A schematic diagram of TALT elements. The upper panel shows the structure of TALT elements. The table summarizes the characteristics of TALT1 and TALT2.

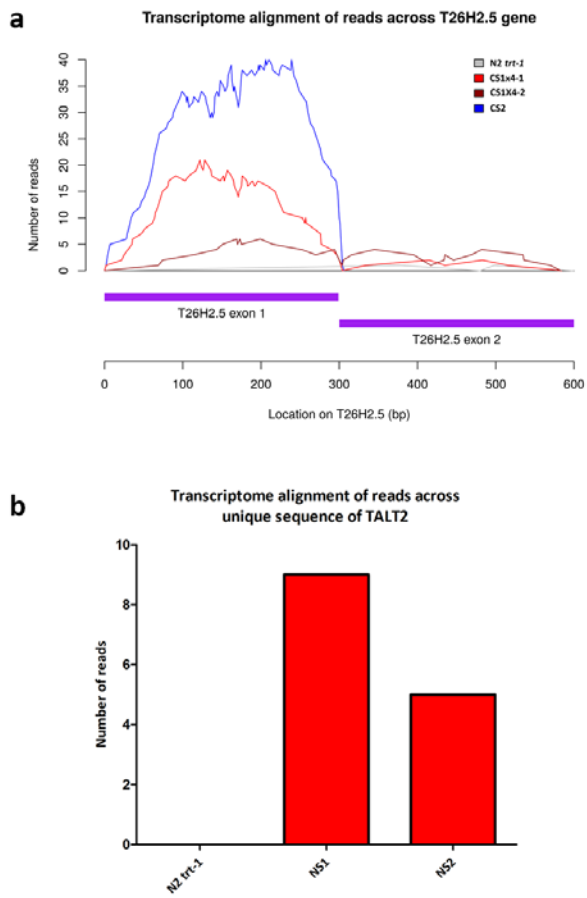


Figure 30. Transcripts of unique sequence of TALT were increased. (a) Reads of mRNA sequencing aligned to specific region of TALT1. Reads of CS1 and CS2 transcript specifically aligned to exon 1 of T26H2.5. **(b)** Reads of mRNA sequencing aligned to specific region of TALT2. Reads of NS1 and NS2 transcript specifically aligned to left proximal telomere of chromosome I. In collaboration with M. Hills.

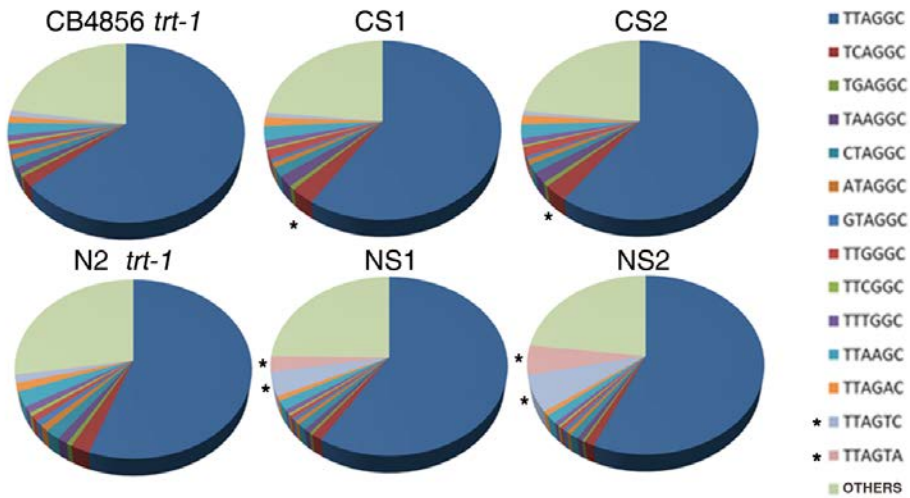


Figure 31. Variant repeats were increased in telomerase-independent survivors. Differences in variant telomere repeat usage were detected in different survivors. Telomere reads were defined as sequence reads with at least 6 non-consecutive TNAGGC repeats. The frequency of each variant within these reads was calculated and used to create pie charts. Asterisk (*) represents increased variants repeats in telomerase-independent survivors. In collaboration with M. Hills.

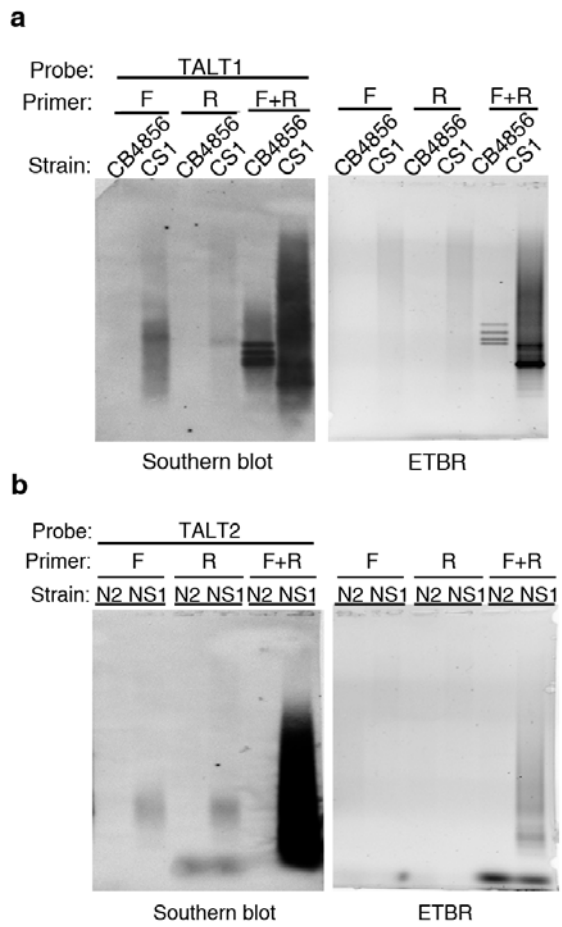


Figure 32. TALT show head-to-tail repetitive patterns. (a) Amplified TALT1 exhibits tandem repeat patterns. PCR was done with single primer, either forward or reverse, to detect head-to-head or tail-to-tail orientation. To detect tandem repeat, both forward and reverse primer was included in PCR reaction. Amplicons were separated by gel electrophoresis and hybridized with TALT1 probe. **(b)** Amplified TALT2 exhibits tandem repeat patterns.

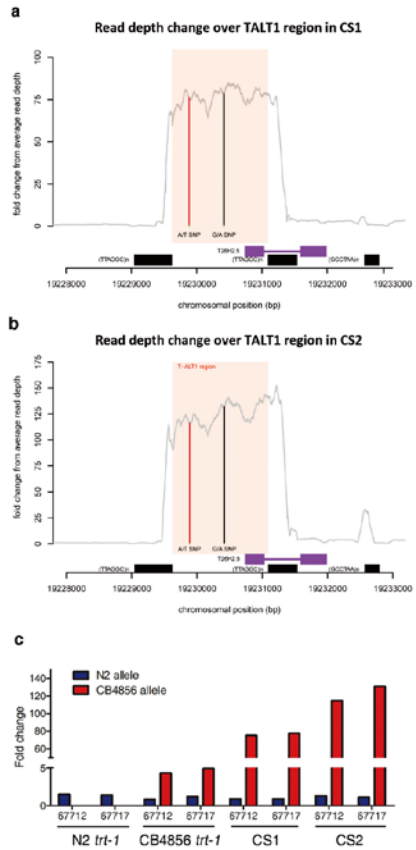


Figure 33. Two SNPs exist in TALT1 locus. **(a)** Read depth change over TALT1 region in CS1 **(b)** Read depth change over TALT1 region in CS2 **(c)** Fold change from average coverage of SNVs (WBvar00067712 and WBvar00067717) from WGS data of N2 *trt-1*, CB4856 *trt-1*, CS1 and CS2. N2-associated haplotype (blue) and CB4856-associated haplotype (red) are plotted separately. Due to large difference of value, Y-axis is segmented. In collaboration with M. Hills.

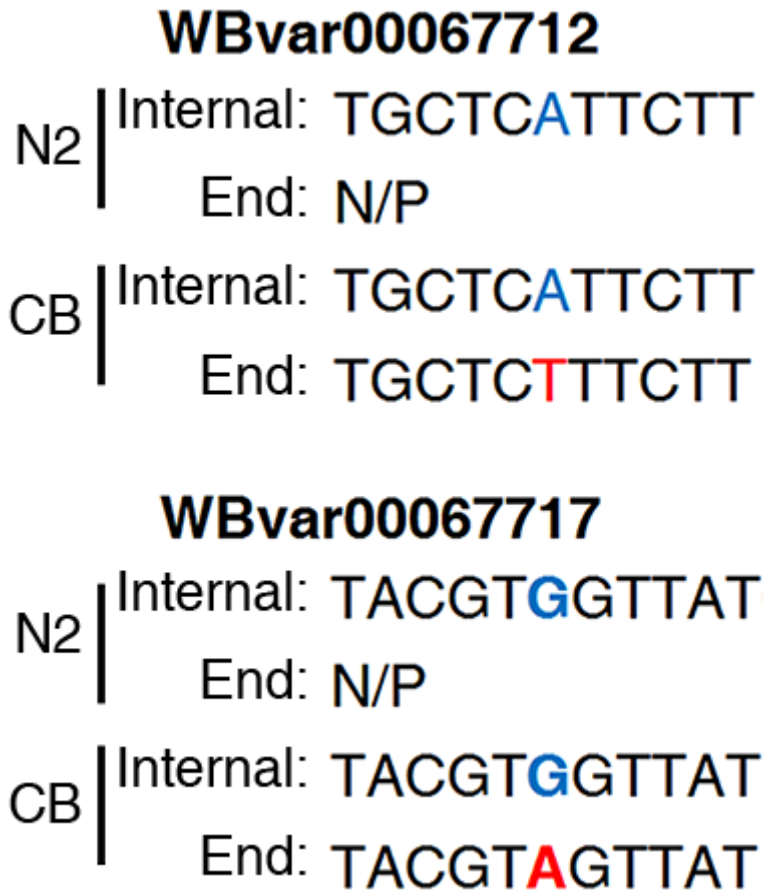


Figure 34. SNPs of CB4856 were located in telomeric TALT1. Nucleotide change of WBvar00067712 and WBvar00067717 in internal and telomeric TALT1. N2-associated alleles are blue while CB4856-associated alleles are red.

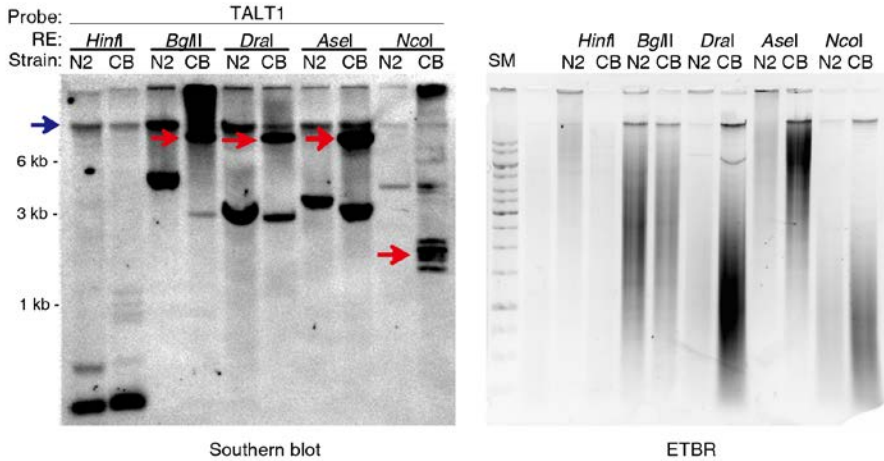


Figure 35. TALT locus have distinct restriction fragments length polymorphism. CB4856 TALT1 shows distinct restriction fragments length polymorphism (RFLP) patterns compared to N2. Fragments that do not exist in N2 was detected in CB4856. Using enzymes that did not cut TALT1, CB4856 TALT1 had other large products that were unexpected from sequence information. The blot was probed with TALT1. Blue arrow indicate incompletely digested fragment of genomic DNA. Red arrows indicate another copy of T26H2.5 locus.

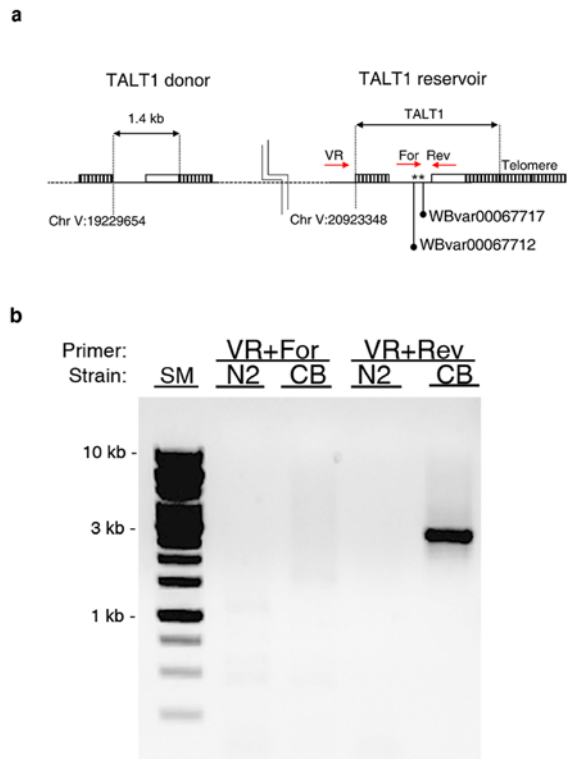


Figure 36. *Trans*-duplication of TALT1 reservoir to other chromosomal ends in CB4856 backgrounds. (a) 'CB4856-derived' SNVs were located at the right end of chromosome V in CB4856, but not in N2. Asterisks indicate CB4856 SNVs. Dashed line indicates genomic position. (b) Breakpoint PCR was performed with primer pairs (red arrows in a) designed to anneal to the proximal telomere region of chromosome V (VR) and TALT1 in forward (For in a) or reverse (Rev in a) orientation. CB, CB4856.

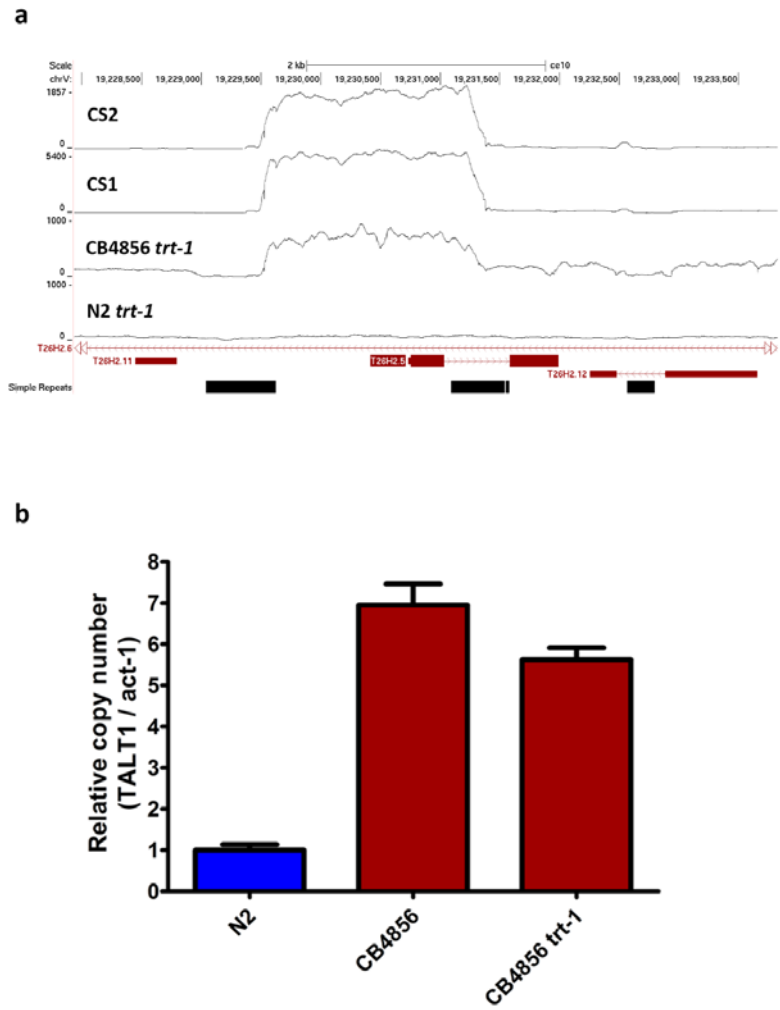


Figure 37. *Cis*-duplication of TALT1 in CB4856. (a) Reads of TALT1 is higher in CB4856 *trt-1* than N2 *trt-1* **(b)** Copy number of TALT1 DNA is higher in CB4856 than N2. qPCR using TALT1 specific primers show 8 folds increase in CB4856 than N2. Fold changes are normalized with single copy gene, *act-1*. *n* = 3. Bar is ± S.D.

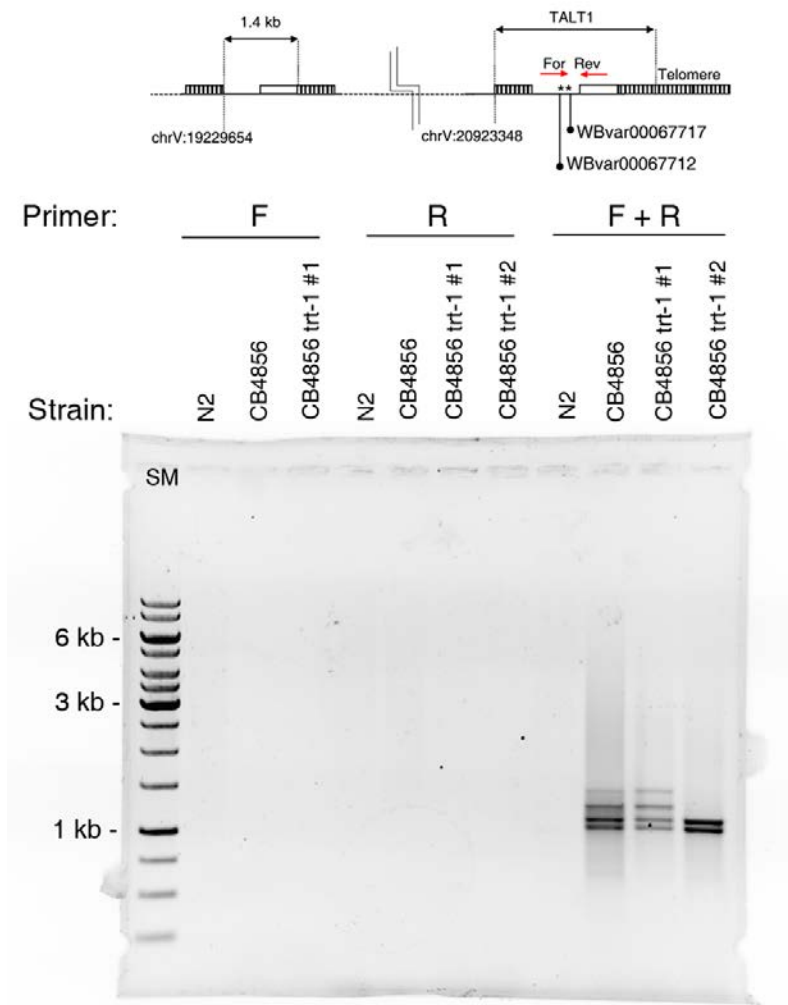


Figure 38. TALT show head-to-tail repetitive patterns in CB4856. TALT1 elements were tandemly duplicated in CB4856. F and R primers were specific for the CB4856-derived allele. TALT1 amplification was detected by southern blot with a TALT1 probe. Blue arrow: uncut DNA. Red arrow: CB4856-specific restriction fragment.

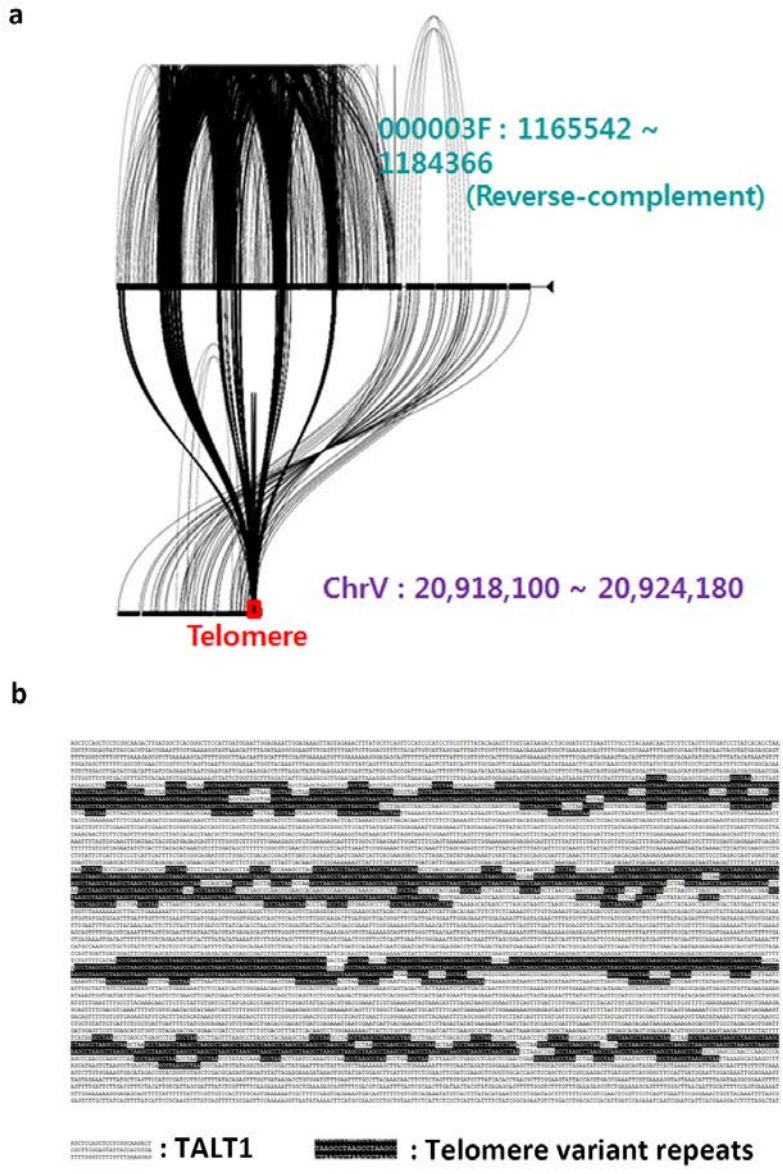


Figure 39. CB4856 have multi-copies of TALT1. (a) Multi-copies of TALT1 was constructed in single PacBio contig of chromosome V. **(b)** Sequence of right arm of chromosome V in PacBio contig.

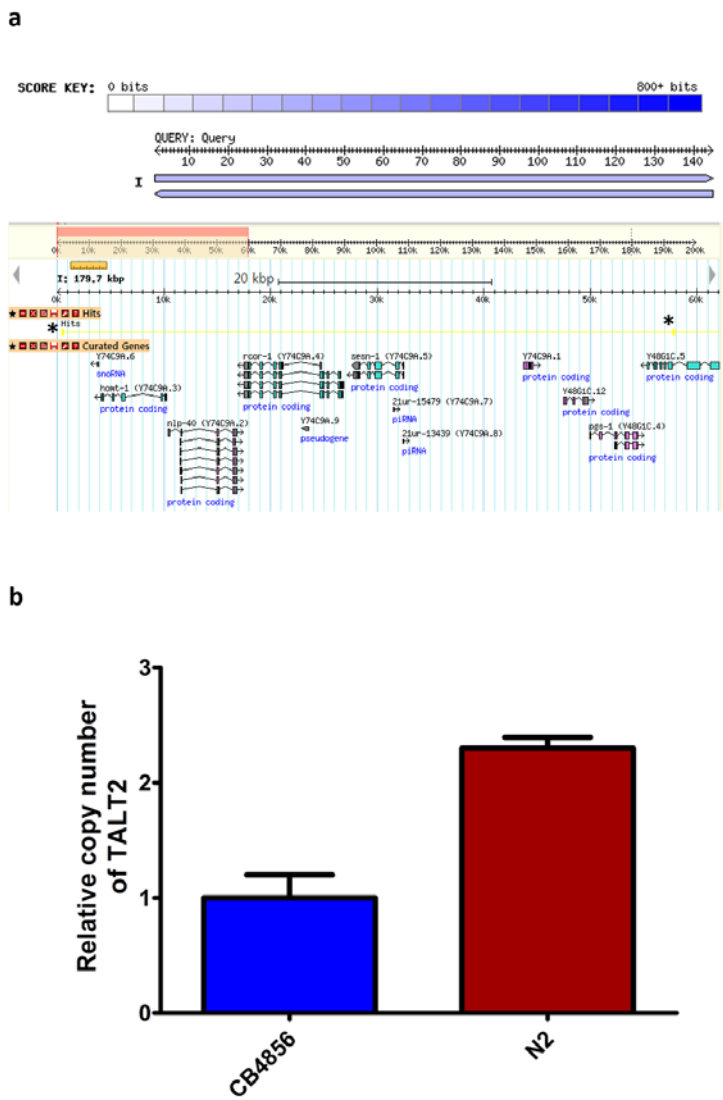


Figure 40. N2 have multi-copies of TALT2. (a) BLAST result of unique sequence of TALT2. TALT2 exist two copy in chromosome I **(b)** Copy number of TALT2 specifically increased in N2 assessed by qPCR and normalized with *act-1*. *n* = 3. Bar is ± S.D.

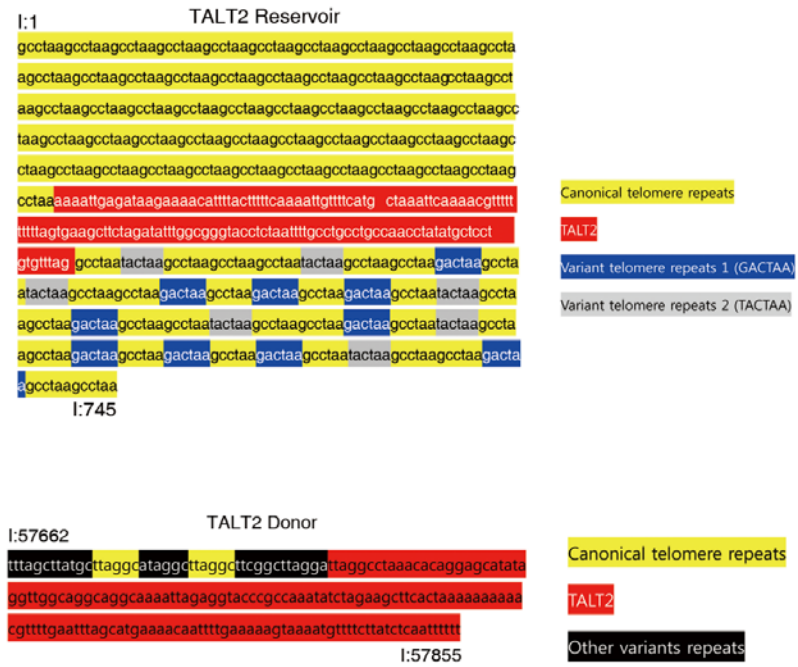


Figure 41. Flanking sequences of TALT2 in telomere-adjacent region and internal were different from each other. Sequence information of TALT2 donor and reservoir site. Red: TALT2, Yellow: canonical telomere repeats, Blue: variants telomere repeats 1 (GACTAA), Gray: variant telomere repeat 2, Black: other variants repeats.

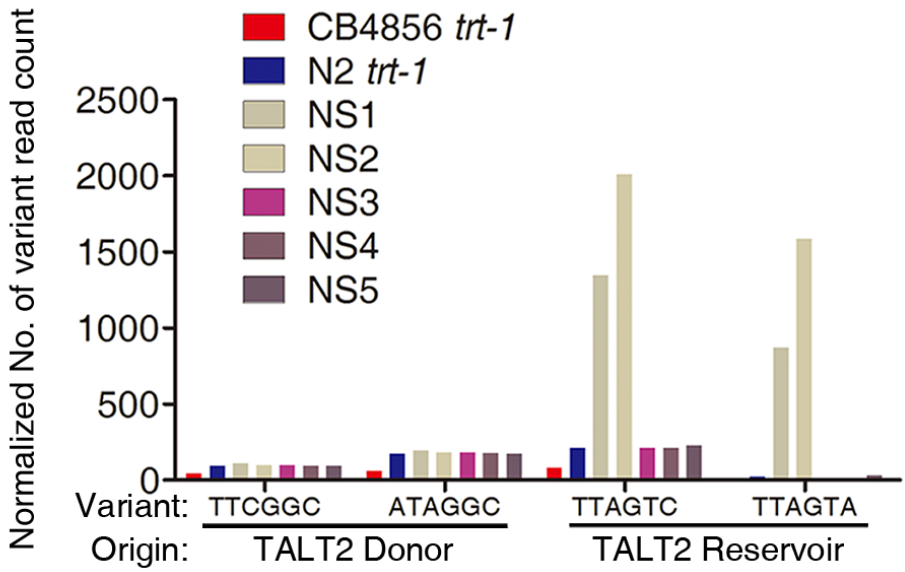


Figure 42. *Trans*-duplication of TALT2 reservoir to other chromosomal ends in N2 backgrounds. Normalized number of variant read count from WGS data of CB4856 *trt-1*, N2 *trt-1* and all the NS survivors. Variant read count was normalized to total sequencing coverage of each sample. In collaboration with M. Hills.

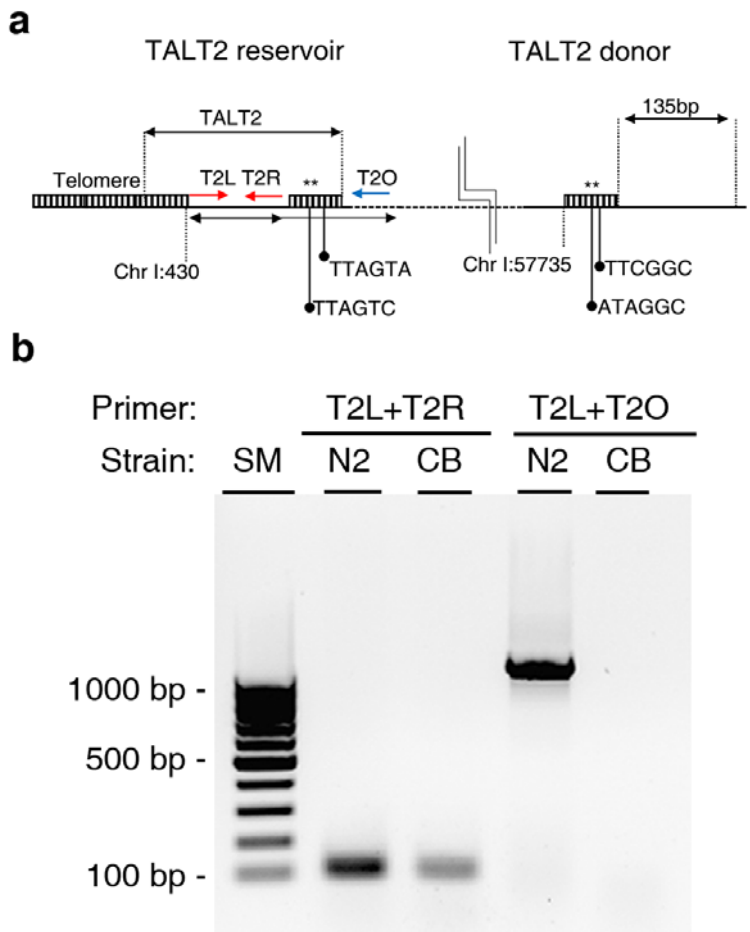


Figure 43. TALT2 reservoir only exist in N2 backgrounds. (a) Diagram of TALT2 in N2 genome. Primers used in **b** are indicated as red arrows. Specific variant repeats (asterisks) of TALT donor and reservoir are indicated. **(b)** TALT2 reservoir is absent in CB4856 wild isolate. PCR amplicons from indicated primers were fractionated by gel electrophoresis and stained with ethidium bromide. N2, N2; CB, CB4856.

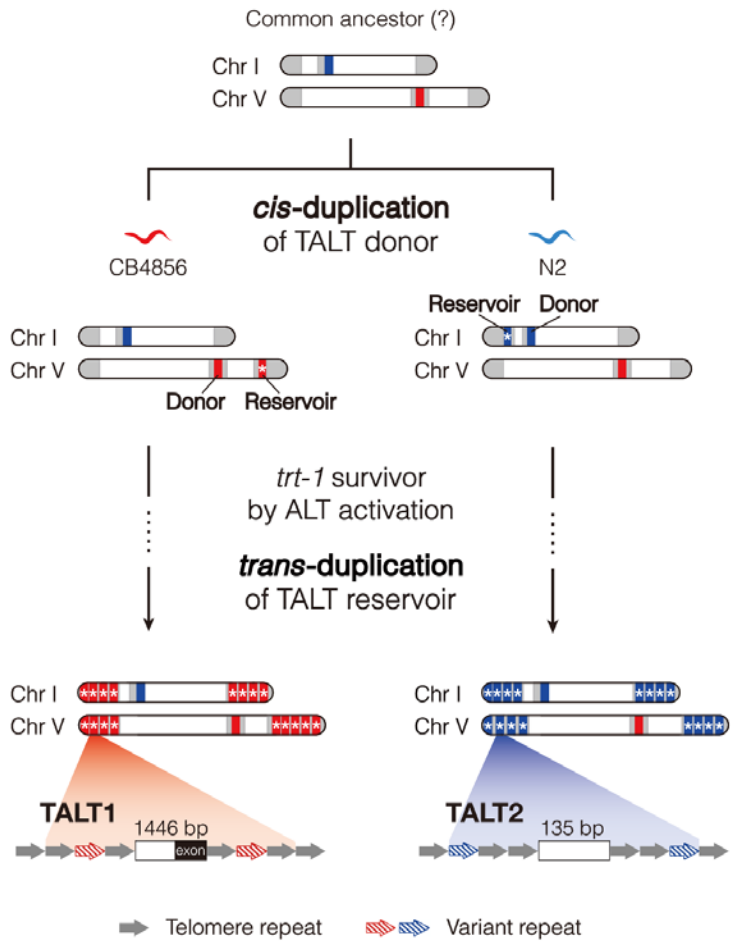


Figure 44. A model for TALT-mediated ALT. Internal genomic regions with ITS are used as template of ALT in *C. elegans* telomerase-independent survivors. The model illustrates that *cis*-duplication of TALT donor into telomere occurred independently in nature without telomerase loss. The TALT reservoir at a proximal telomere region is used as a template for telomere maintenance at the time when telomere crisis emerging. In collaboration with B. Seo and D. S. Lim.

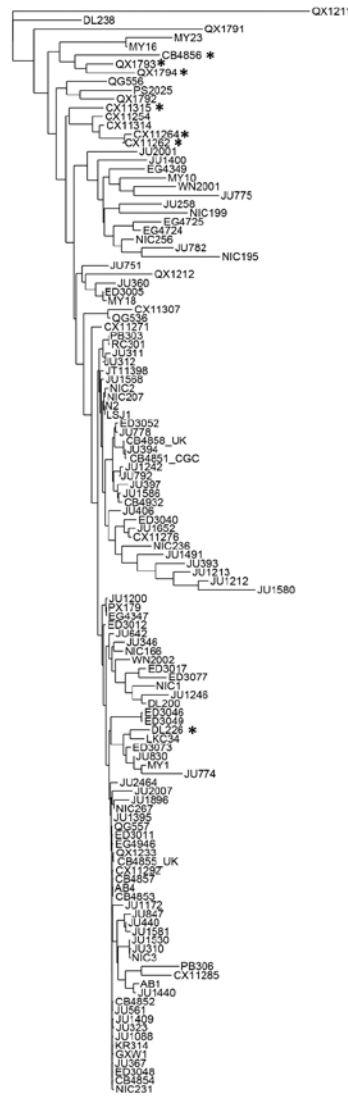


Figure 45. *Cis*-duplication of TALT1 to proximal telomere has occurred independently multiple times. The phylogenic tree of 152 natural isolates. The asterisk strains contain CB4856-type TALT1 on chromosome V. In collaboration with D. Cook and A. Erik.

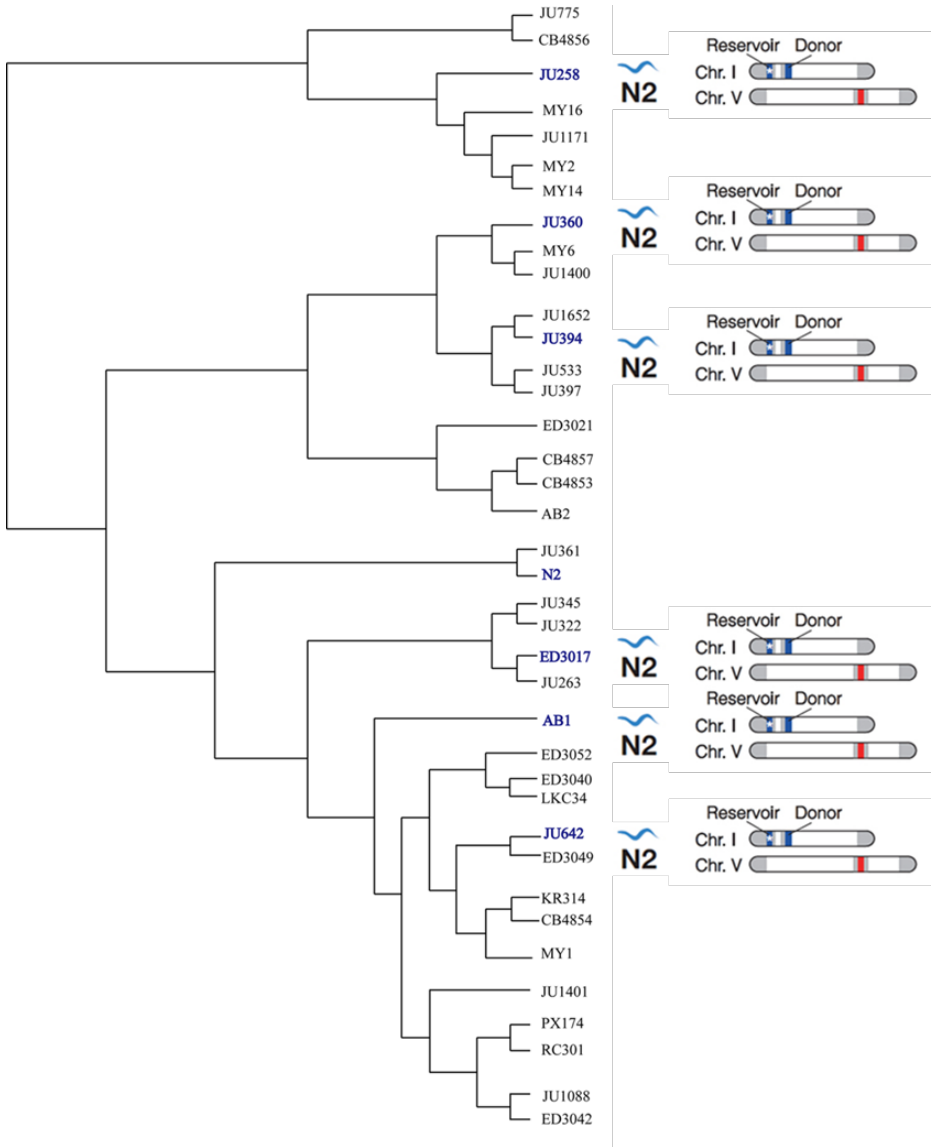


Figure 46. *Cis*-duplication of TALT2 to proximal telomere has occurred independently multiple times. The phylogenic tree of 38 natural isolates. The blue color-coded strains contain N2-type TALT2 on chromosome I.

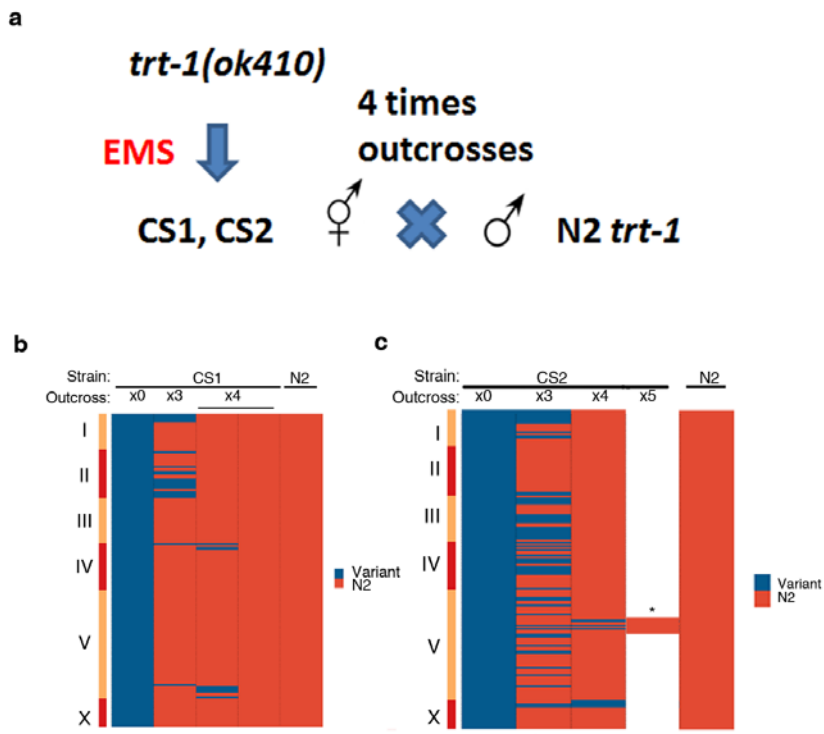


Figure 47. No single mutation was responsible for inducing CS survivor. (a) Schematic image of CS outcross **(b)** Mutational heatmap of CS1 across the genome. None of the variants that were present in the initial isolates were maintained after outcrosses. **(c)** Mutational heatmap of CS2 across the genome. CS2 survivor was outcrossed with N2 *trt-1* by five rounds. Remaining variant on X chromosome was affecting pseudogene (Y35H6.3). Asterisk indicates N2 genetic background confirmed by snip-SNPs mapping using pKP5113, pKP5114 and pKP5116 SNPs. In collaboration with H. Oh and M. Choi.

Strain	ORF name	Gene name	Status	Predicted function	Southern lane
CS1x4-1	2L52.1		deletion	ZF TF	1
CS1x4-1	2RSSE.1		deletion	RhoGAP domain	9
CS1x4-1	2RSSE.2		deletion	RhoGAP domain	10
CS1x4-1	Y67D2.6		deletion	RNA helicase	11
CS1x4-1	Y38C1AB.4	frm-5.2	deletion	FERM domain	13
CS1x4-1	Y38C1AB.8	frm-5.1	deletion	FERM domain	13
CS1x4-2	Y66A7A.5	ceh-91	deletion	reproduction 관련	4
CS1x4-2	Y66A7A.4		deletion	unknown	3
CS1x4-2	Y38C1AB.4	frm-5.2	deletion	FERM domain	13
CS1x4-2	Y38C1AB.8	frm-5.1	deletion	FERM domain	13
CS2x4	Y74C9A.3	homt-1	deletion	Hydroxyindole-O-MethylTransferase homolog	7
CS2x4	W02B12.11		deletion	carbohydroxy	2
CS2x4	K04F1.15	alh-2	deletion	dehydrogenase	5
CS2x4	B0213.10	cyp-34A5	deletion	cytochrom P450	8
CS2x4	Y38A10A.7		deletion	unknown	12
CS2x4	cTel55X.1		deletion	unknown	6

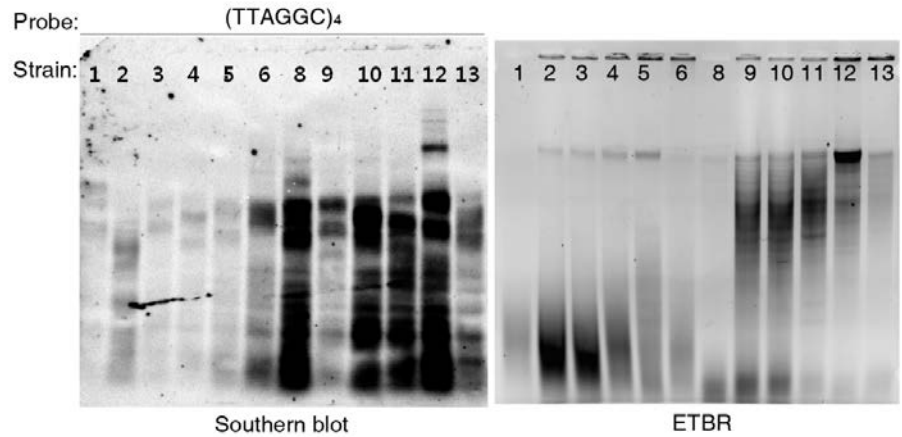


Figure 48. No single deletion mutation was responsible for inducing CS survivor. The upper panel show list of candidate genes that have deletions in exons by comparative genome hybridizations. The bottom panel show knockdown of these genes cannot induce CS survivors and no telomere lengthening detected. This blot is hybridized with DIG-TTAGGC*4.

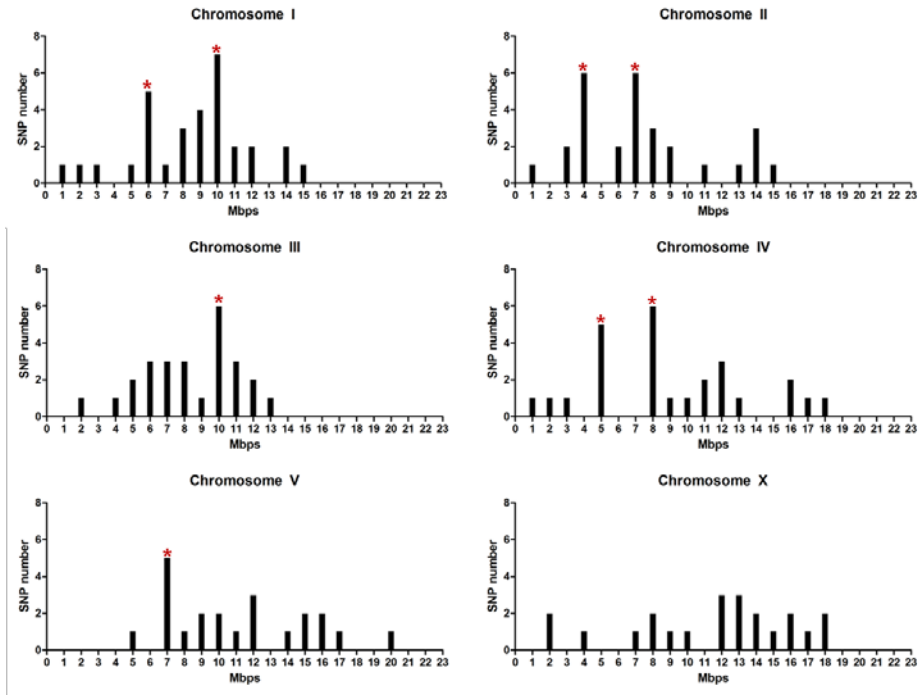


Figure 49. Single mutation linkage was not shown in NS1 survivor. This result show mutation linkage mapping by WGS in NS1 survivor. In collaboration with E. Kim.

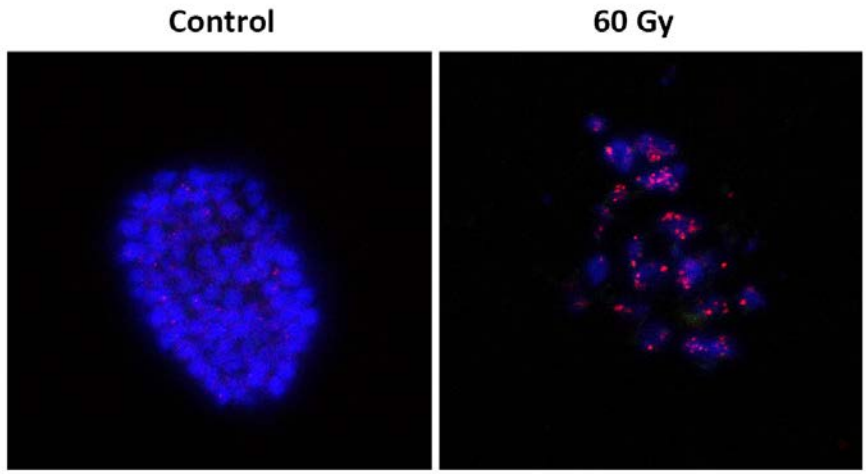


Figure 50. NS survivors were induced after gamma irradiation of N2 *trt-1* mutant. FISH analysis for telomere length of gamma-irradiated survivor. Telomere was detected by Cy3-TTAGGC*3 PNA probe (red) in the embryo. DNA was counterstained with DAPI (blue). Control was established by massive chunking. 60 Gy represent gamma survivor established by gamma irradiation.

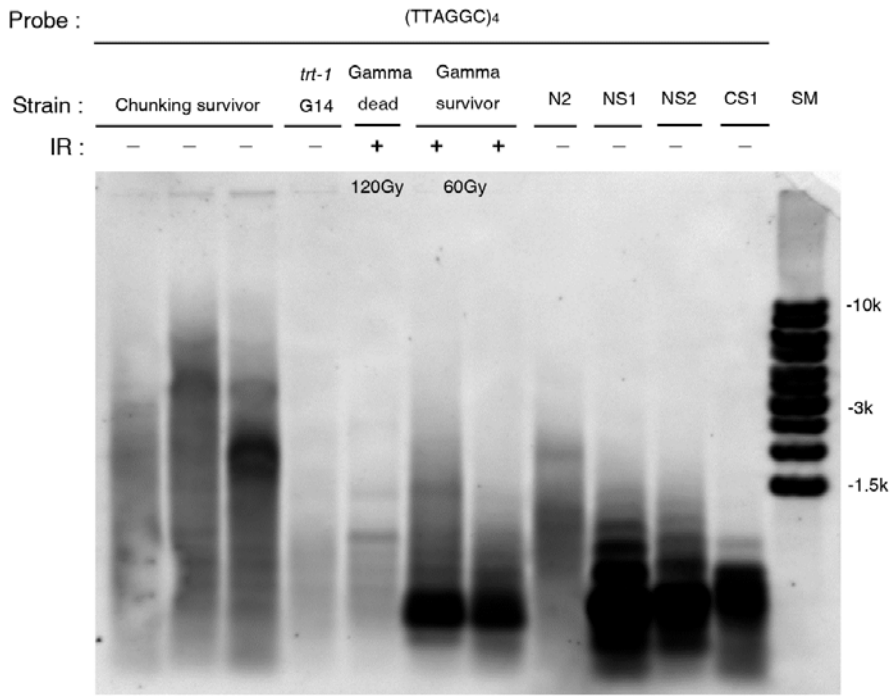


Figure 51. NS survivors were induced after gamma irradiation of N2 *trt-1* mutant. After treating 60 Gy gamma irradiation, we isolates two independent survivors from 30 gamma irradiation-treated plate of N2 *trt-1* worms. NS1 and NS2 were used as positive controls as they have TALT2 insertion. This blot is hybridized with DIG-TTAGGC*4.

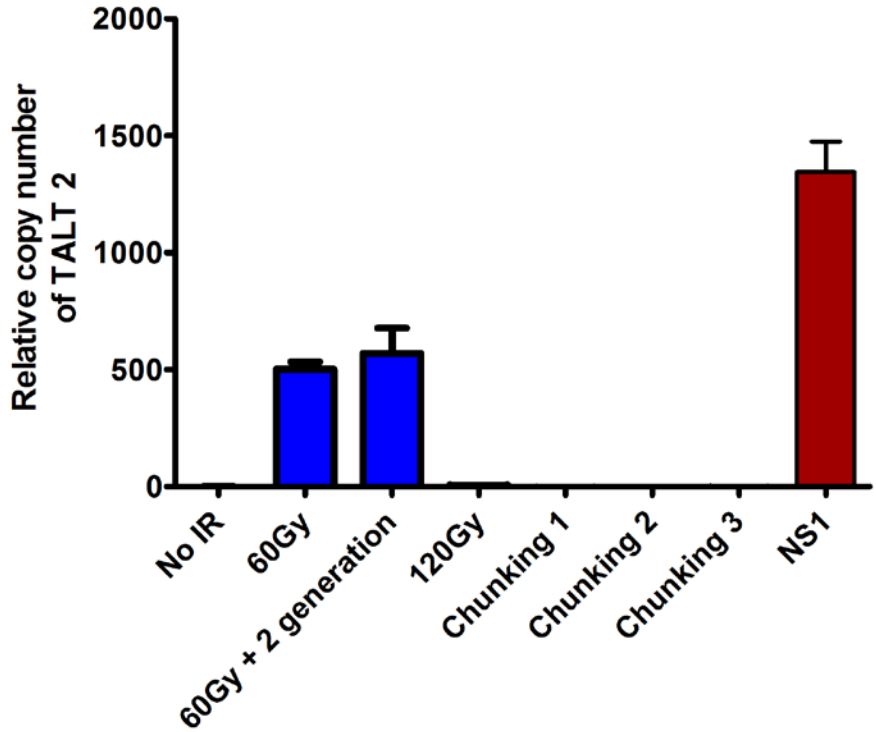


Figure 52. TALT2 was increased in gamma-irradiated N2 *trt-1*. The figure show qPCR of ALT survivors from gamma irradiation (labeled as 60 Gy and 60 Gy +2, respectively) for detecting TALT2 amplification. NS1 survivor was used as a positive control, and worms maintained by chunking were used as negative controls.

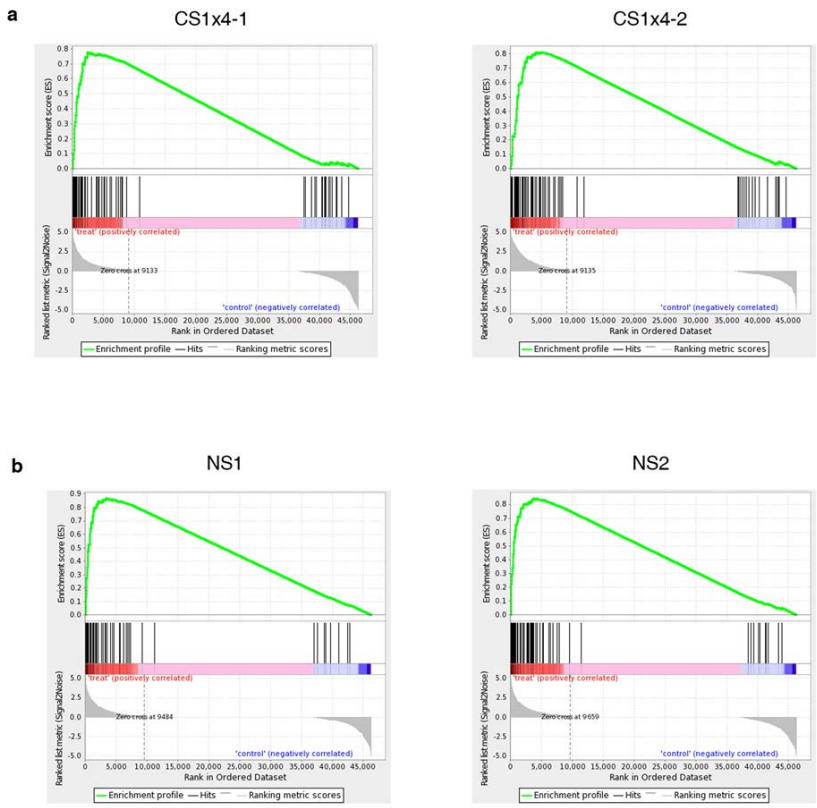


Figure 53. Telomerase-independent survivors showed gene set by induced gamma-irradiation was enriched. (a) GSEA results of CB4856 survivors. Gamma-ray responsive gene sets are enriched in CS1x4-1 and Cs1x4-2 survivors. This gene set significantly enriched at FDR *q*-value < 0.005 in both strains. **(b)** GSEA results of N2 survivors. Gamma-ray responsive gene sets are enriched in NS1 and NS2 survivors. This gene set significantly enriched at FDR *q*-value < 0.005 in both strains. In collaboration with H. Oh.

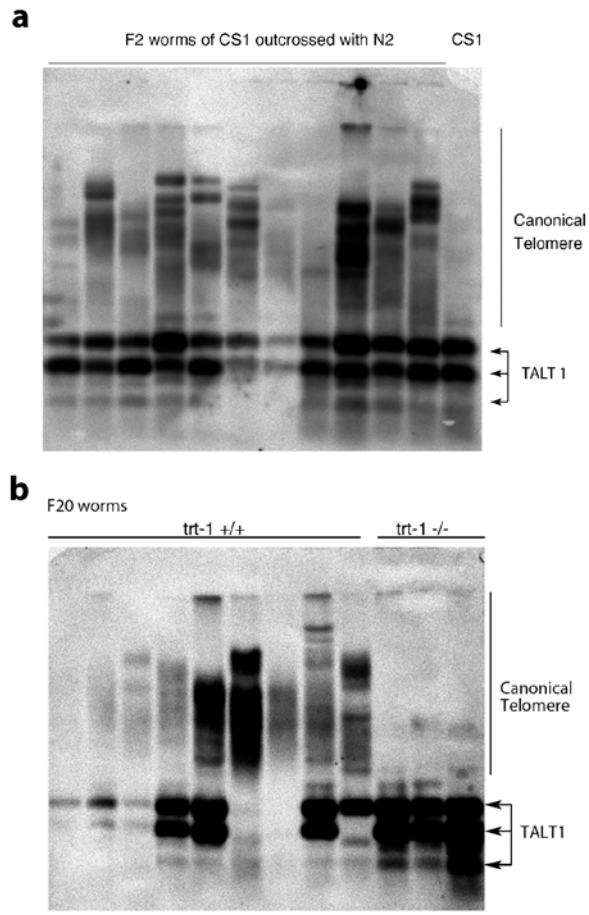


Figure 54. TALT1 prevailed many generations after backcross with wild-type N2. (a) All F2 animals after outcross contained both TALT 1 and canonical telomeres**.** This blot hybridized with DIG-TTAGGC*4. **(b)** After 20 generations, all survivors of the *trt-1* mutant background contained TALT1 telomeres alone. Interestingly, TALT1 telomeres still remained even in the telomerase-positive background. This blot hybridized with DIG-TTAGGC*4.

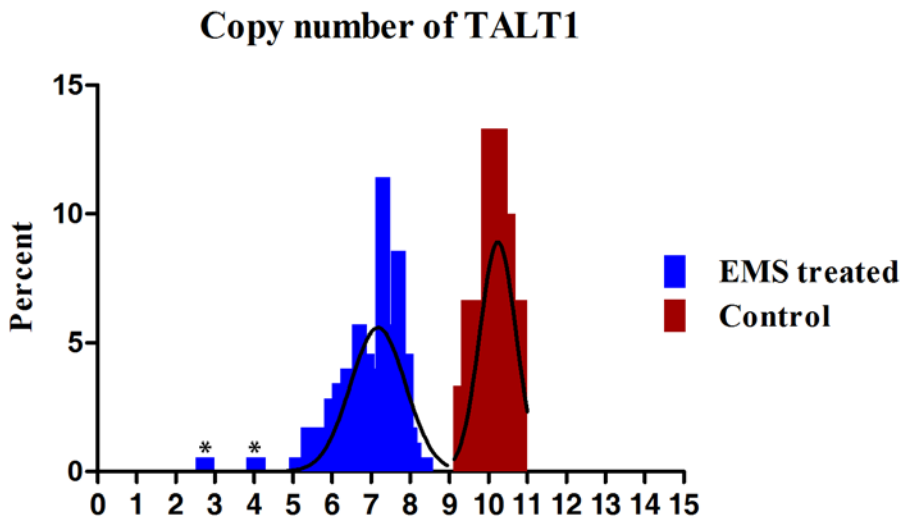


Figure 55. TALT1 levels decreased in 4 EMS-treated CS1. After 5 generating of EMS-treating, 200 worms showed sterile phenotype. qPCR show 4 worms had decreased levels of TALT1. Interestingly, all EMS treated worms had lower TALT1 levels than control.

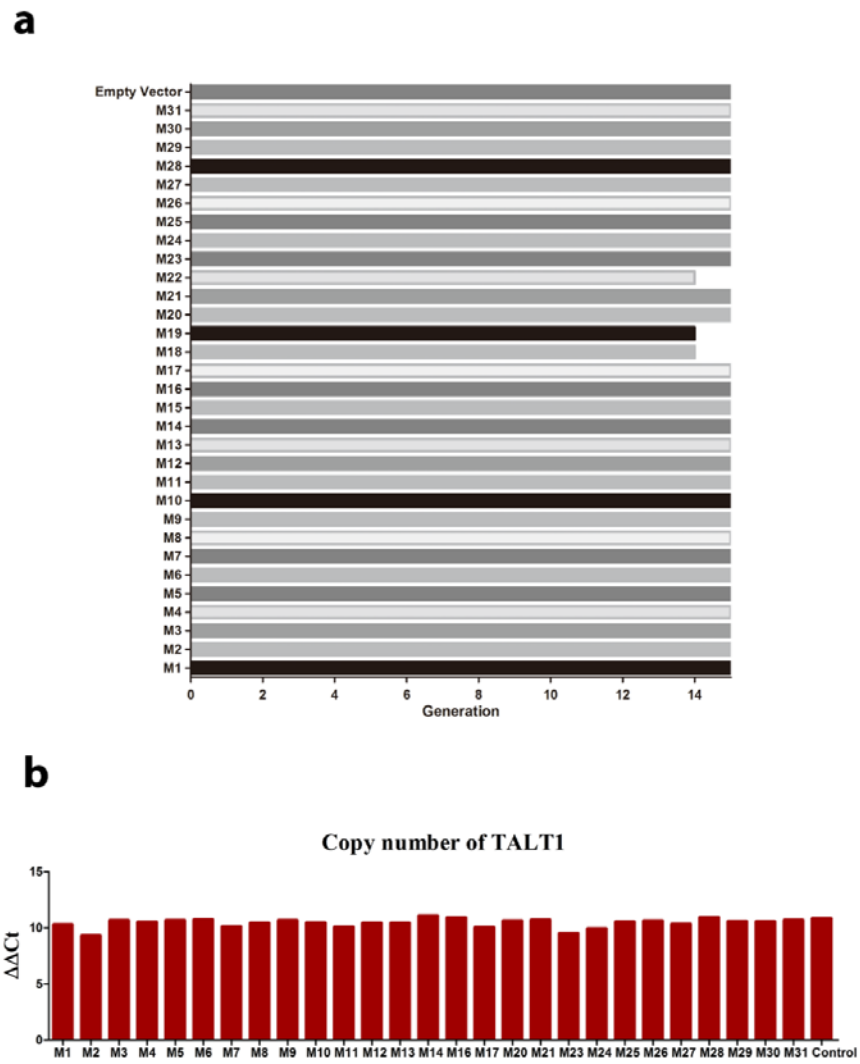


Figure 56. Meiotic regulators were not required for TALT1 duplication. (a) 31 candidate genes suppressed in CS1 by feeding RNAi. 3 genes showed sterile phenotype but these genes also induced sterile in wild type. **(b)** Copy number of TALT1 in CS1 survivor didn't changed after RNAi.

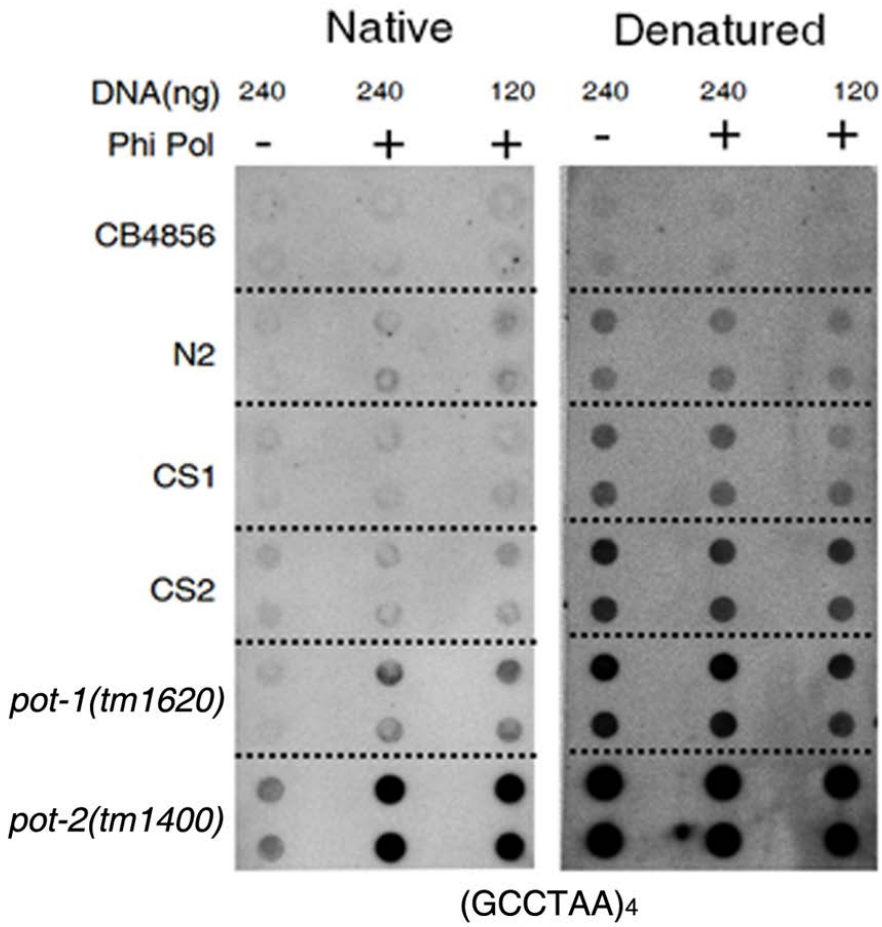


Figure 57. C-circles were not detected in ALT survivors. Left panel shows native condition that detects only single strand DNA amplified by phi polymerase. As a negative control, the same blot was denatured and double stranded template DNA was detected in right panel. *pot-1(tm1620)* and *pot-2(tm1400)* was used as positive control which is known to have C-circles.

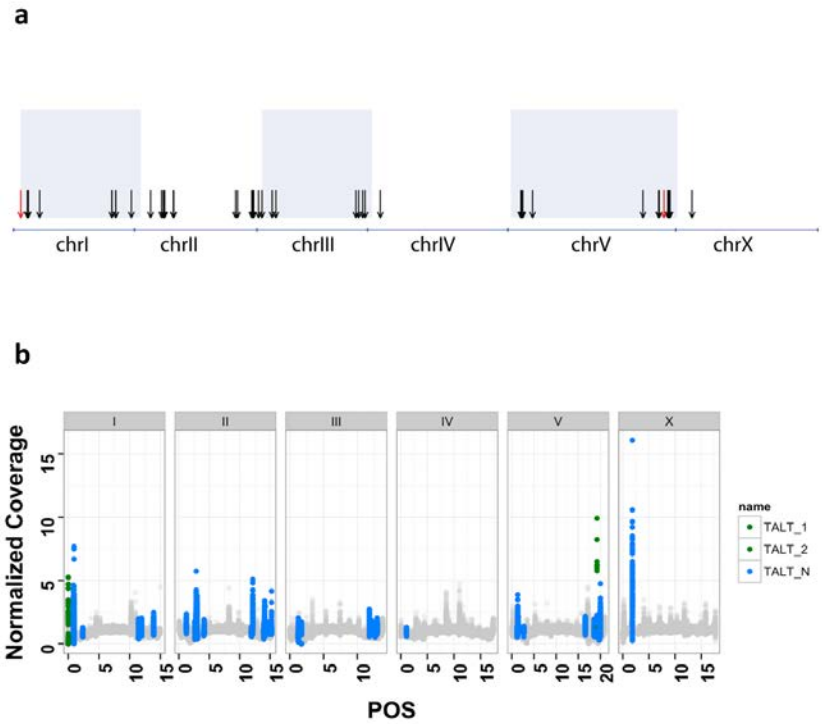


Figure 58. Potential TALT candidates in natural isolates. (a) Potential TALT candidates. Genomic locations of potential TALT are marked by arrows. Red arrows indicate TALTs identified in this study. Potential TALT elements are defined as a region between 100 - 2,000 bp flanked by ITS. (b) Distribution of depth of coverage in TALT regions across 152 wild isolates. TALT regions were plotted behind normalized genome-wide coverage in 100kb bins. TALT regions that rise above normalized coverage suggests that they are highly specific to the repetitive region. In collaboration with M. Hills, D. Cook and A. Erik.

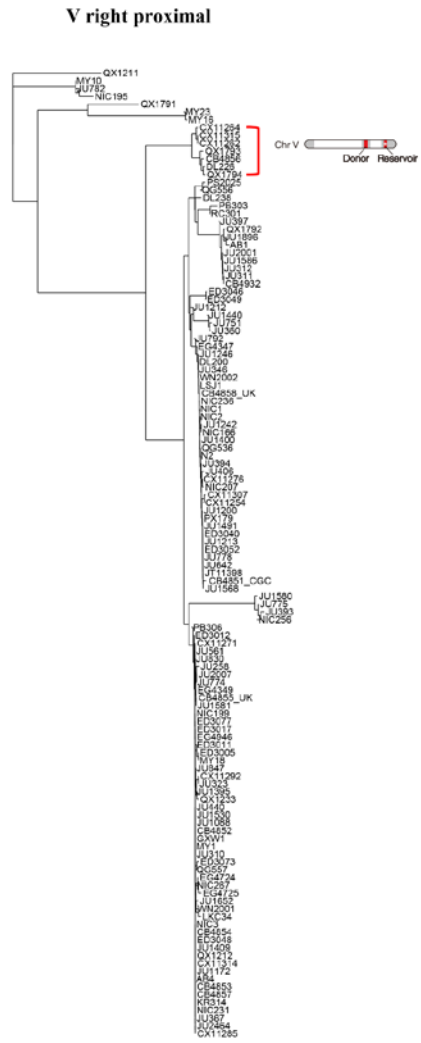


Figure 59. Strains having amplified TALT1 were tied closely. This Phylogeny tree was drawn by SNP within 25kb from telomere of chromosome V right. 7 strains (CX11264, CX11315, CX11262, QX1793, CB4856, DL226 and QX1794) from 152 natural isolates have high level of TALT1 copy number. In collaboration with D. Cook and A. Erik.

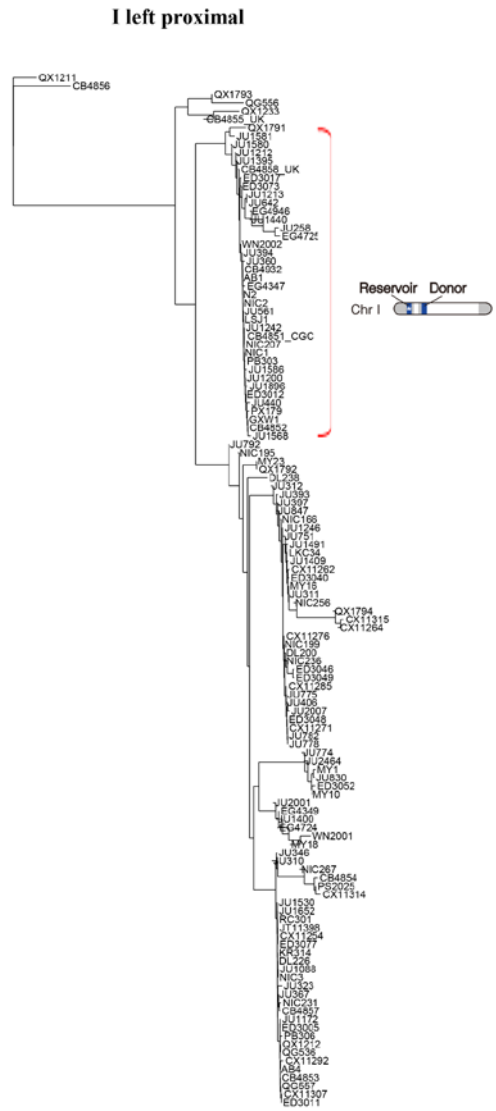


Figure 60. Strains having amplified TALT2 were tied closely. This Phylogeny tree was drawn by SNP within 25kb from telomere of chromosome I left. 40 strains (Table 6) from 152 natural isolates have high level of TALT1 copy number. In collaboration with D. Cook and A. Erik.

Table 1. The list of contigs constructed using telomere-containing reads.

Chr.	Location	Sequence*	PCR
IR	15071743 ~ 15072072	GGCACTAGTACTATGCGCCGCGAGACACACACTACCATCA CCAACACAGCCCCCTGAGACATCACGTCTCAGACCAACTG TCACCCCCCTGACACATATAGGAGTGTGCGGGGAGGGAGT ATTATAAAAACCACGGAAAGCCGGCAGTGTGCAATTTGAG AGACGGCAGACAACGCGCGCACTACCCCCCACCACACCGA CCCCCTTCCTCGAGACACCGTTTTTCTCAGCCGTCTCTCTG GAGACCTCGGCCGCTTGCACAAAATTTTCTATCATTTTGAG ACTTTGGTGGTGTGTCCCCCTCTCCACCAGCACTGCGGCTG CTGCCTACAGCAGAGTGGAGAATACGTGGGGGATAAGTAA GATTTTTCTATTTTTCATAAGCCTA	O
IIR	15255618 ~ 15256023	CGAACTTCGAGAACAAACTTAGGCTTATTACGGCCGACGA TTTCAAATTAAAAAAAAAACTGAAAAATTCAGGGCGAATT TGTTAATATTTTTATGTTAGCAACCATGAGAGACGCAGAG AGAGAGACAGAAAGATGAAAAAGGGCGCTTGCGGACGTC GTCGGGTAGGGTGCCTGAAGTCGATGCGAAGAGATGGGA GTGGAGAGAGTGTGGTCACCATGAGAGACGCAGACATGTA CACACATACTCTACGTCTCTCTGAGAAGAGCCGCTGTGCTA CTACTATGGGTGCAATAGTATGGAGAAGATGAGAGATGAA CCTTTTTAGATTAGAATATTTTGAAAACGTGGGATTCGTTT TTAAATAGATTTTACTGAAATATTGCTGCGAAATTTCCATC AATTTTTTTTCTTTTTTGGAAAAACACAAAAATCAAACTAC GGTATAAGGTTGCTAACATAAAAATATTAACAAATTCGCC CTGAATTTTTCAGTTTTTTTTTTAATTTGAAATCGTCGGCCG TAATAAGCCTAAG	O

IIIL	119 ~ 332	ATTTATCGATTCTTATCGATTTTCTTCTCTTTCCGAACCTTT TCGGAATCAAAAAGCCGTGAATCCATAGATTCCGTGCTTT CTCAGACTTTTCGAGGCCTAATTTTGGTCGAAAAGCCCGAT TTTATTTTTATCTGATTTCCGACTCTTTTCGGACTCAAAAAG TTATAAATCCTTGAAATTTCATAGGATTTCGCGCTT**TTAGG** **CTTAGGCT**	O
VR	20923658 ~ 20923995	AACGGACAAAAACTTTAAAAACCTCCTGTAAAAGAACCTG GGCTGAGCAAAAGCACATGAAAATTTTTTGAATAAAATAT CATACTTCATACACCATATACCCATAAAGTAATAAGAATC TAATGGCAACAGAGGATACTGGTAAGAATCTAATGGCGAC AGAGGATACTGGGAATTTTTATCTGTAACTTATTGGCATCC TTAATAGCATAAAGGTTTATTGAATGTTTTCAAGCTATTTT GAAAATTCTTTAATTCCGCGAAGTTTCTTGATTTCTAAATT ACTGAACAAAAACCAGCTACTTGCGTAAGGTACTGAACCA CTTTTCTGGGGATTTGC**TTAGGCTTAGGCTTAGGCTTAG** **GCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTAAGGC** **TTAGGCACAGGCTCAGGCTTAG**	X
XR	17718698 ~ 17718815	CCAAAGAACAAAAAAAGAAATTAAAATATTTATTTTGCTG TGGTTTTTGATGTGTGTTTTTTATAATGATTTTTGATGTGAC CAATTGTACTTTTCCTTTAAATGAAATGTAATCTTAAATGT ATTTCCGACGAATTCGAGGCCTGAAAAGTGTGACGCCATT CGTATTTGATTTGGGTTTACTATCGAATAATGAGAATTTTC AGGC**TTAGGCTTA**G	X

* Red, TALT sequence; Yellow, Telomere repeat

Table 2. The list of junctions between chromosome end and TALT in ALT survivors.

Strain	Chr.	Sequence*
CS1	IR	TCCAACCCGAATGATGGAGATTTTTTAGGTAGCTTTTTTTAGACCAATA GCAAATTCATAGTCCGACTAGCCTATAGACTTGGGCTTAGGCTTAGGC TCAGGCTCAGGCTTAGACTTATGCTAAGGCTTATGCTTTTAATTAGGCT TAGGCTTAGGCTAAGGCATTCACTTAGGCTTAGGTTTAGGCTTAGGCTT CGGTTCGAGCTGAGGCTCAGACTTAAGCTTAGGCTCAGGCTTAGGCTT AGGCTTAGACTTTGGATTAAGCTTAGGCTTTGGTATAGGCTTAGGCTTA GGTTTAGGCTTAGGCTTGGGCTTGGGTTGGGCTTGGGCTTGGGTTGGGC TCAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCT TAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTT AGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTA GGCTTAGGCTTAGGCTTAGGCTTATGAAAAATAGAAAAATCTTACTTA TCCCCCACGTATTCTCCACTCTGCTGTAGGCAGCAGCCGCAGTGCTGGT GGAGAGGGGGACACACCACCAAAGTCTCAAAATGATAGAAAATTTTG TGCAAGCGGCCGAGGTCTCCAGAGAGACGGCTGAGAAAAACGGTGTC TCGAGGAACGGGGTCGGTGTGGTGGGGGGTAGTGCGCGCGTTGTCTGC CGTCTCTCAAATTGCACACTGCCGGCTTTCCG
CS1	IIR	GTCTTATCACCAAACTCTGTATAAAACGAAGGATGGGATGGAACTGAA GCATAAAGTTTCTACTAACTTTCTCCAATTTCTCCAATTCCATCAATGG AAGCCCGTGAGCCCTCAAGTCTTGCCGAGGAGCTGGAGCTGGTGCCCC CCGAGCTTCGATCGAACTTCGAGAACAAACTTAGGCTTATTACGGCCG ACGATTTCAAATTAAAAAAAAAACTGAAAAATTCAGGGCGAATTTGTT AATATTTTTATGTTAGCAACCATGAGAGACGCAGAGAGAGAGACAGA AAGATGAAAAAGGGCGCTTGCGGACGTCGTCGGGTAGGGTGCCTGAA GTCGATGCGAAGAGATGGGAGTGGAGAGAGTGTGGTCACCATGAGAG ACGCAGACATGTACACACATACTACCCTTT

CS1	**IIIL**	GGCGAATTTTGCGCTCAGAACGCAGACTCAGTTCAGACTCGAAAGTTA TTAATCCTTGAAATTTCATAGGATTTCGCGCTTTTAGGCTTAGGCTTAG GCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGG CTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGC TTAGGCTTAGGCTTAGGCTTAGGCTTATGAAAAATAGAAAAATCTTAC TTATCCCCCACGTATTCTCCACTCTGCTCAAGAAAAGTTCTTTAGCCAG CTCCTTTAATTGTTCCAGCGCCTTCGAATGATCCAAAGCAACCAGAATA ACTTTTTTCCCAGACTTGTGAATAAACGTCACAGAAACCAGATCGGTTC CGTCGTCTCTACGACCACATGCTC
NS1	**VR**	TAAGCCTAAGACTAAGCCTAATACTAAGCCTAAGCCTAAGACTAAGCC TAAGCCTAATACTAAGCCTAAGCCTAAGACTAAGCCTAATACTAAGCC TAATACTAAGCCTAAGCCTAAGACTAAGCCTAAGCCTAAGCCTAAGCC TAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCAAATCCC CAGAAAGGTGGTTCAGTACCTTACGCAAGTAGCTGGTTTTTGTTCAGTA ATTTAGAAATCAAGAAACTTCGCGGAATTAAAGAATTTTCAAAATAGC TTGAAAACATTCAATAAACCTTTATGCTATTAAGGATGCCAATAAGTT ACAGATAA
NS1	**XR**	CTCGACTCCGTGCACGAGGGCACCGAGGAAAGATAGCCACACGTAGC CGAGGTAGACAATGAACATGCAGGCGGCAATGAACACAAGGCGGAAT GGTGCGGCAAGTTCTCCGAGAATCGGGTACGCGCAGAAGCCAGAAGT GTAGAAGACGTAATGGAGGCTGGAAAATGGTGTGAGGAAGGCAAAAT TAGATGTAGGCGCCAAATATCTAGAGGCTTCACTAAAAAAAAAAACGTT TTGATTTGGCCATGAAAACGATTTTGAAAAAGGAAAATGTTTTGCTAT CGCAATTTTTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAG GCTTAGGCTTAAGCTTAAGCTTAAGCTTAGGCTTAGGCTTAAGGCTTAG GCTTAGGCTTAGGCTTAGCTTAAGCTTAGGCTTAAGTCTTAGGCTTATC TTAGGCTTAGGCTTAGACTTAGGC
NS1	**IIL**	CGACCCGTCATCGTTTACAGGAGCATATAGGTTGGCAGGCAGGCAAAA TTAGAGGTACCCGCCAAATATCTAGAAGCTTCACTAAAAAAAAAACGT TTTGAATTTAGCATGAAAACAATTTTGAAAAAGTAAAATGTTTTCTTAT

CTCAATTTTTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAG
GCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGG
CTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGC
ATAGGCATAGGCATAGGCATAGGCATAGACATAGGCATAGGCATAGG
CATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAG
GCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATA
GGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCAT
AGACATAGGCATAGGCATAGGCATAGGAATAGACATAGGCATAGGCA
TAGGCATAGGCATAGACATAGGCATAGGCATAGGCTTAGGCTTAGGCT
TAGTATTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGTCTT
AGGCTTAGGCTTAGTATTAGGCTTAGGCTTAGTATTAGGCCTAAACAC
AGGAGCATATAGGTTGGCAGG

---

* Blue, proximal telomere region sequences; Red, TALT sequence; Yellow,

Telomere repeat

Table 3. The list of RNAi-subjected genes (DDR and recombination related).

Gene name	Predicted roles	Gene name	Predicted roles
atl-1	DSB response	*lin-35*	RB homolog, tumor suppressor
atm-1	DSB response	*mlh-1*	mismatched repair
brc-1	Recombination	*mre-11*	DSB response
brc-2	fanconi anemia pathway	*msh-2*	mismatched repair
ceh-37	Telomere binding protein	*msh-6*	mismatched repair
cep-1	P53 homolog	*mus-81*	endonuclease
chk-2	Meiotic recombination	*plp-1*	Telomere binding protein
com-1	Meiotic recombination	*pms-2*	mismatched repair
D1081.9	Meiotic recombination	*pot-1*	Telomere binding protein
dna-2	helicase	*prom-1*	Meiotic recombination
dog-1	helicase	*rad-50*	Homologous recombination
drh-3	fanconi anemia pathway	*rad-51*	Homologus recombination
exo-3	AP endonuclease	*rad-54*	Homologous recombination
F55A12.10	Meiotic recombination	*rfs-1*	Homologous recombination
fanci-1	fanconi anemia pathway	*rpa-1*	RPA, replication
fcd-2	fanconi anemia pathway	*rpn-1*	protease
hel-308	helicase	*rtel-1*	helicase
him-17	DSB formation	*set-11*	Histone methyltransferase

Gene name	Predicted roles	Gene name	Predicted roles
him-18	Homologous recombination	*set-25*	Histone methyltransferase
him-3	Meiotic recombination	*spo-11*	DSB formation
him-6	BLM helicase	*top-3*	Topoisomerase IIIa
him-8	Meiotic recombination	*vhp-1*	MAPK
hmg-5	Telomere binding protein	*wrn-1*	WRN helicase
hpl-2	Heterochromatin protein 1	*xpf-1*	ERCC1/XPF endonuclease
hrp-1	Telomere binding protein	*Y39B6A.16*	Meiotic recombination
ku-80	Non-homologous end joining	*zhp-3*	Meiotic recombination

Table 4. The list of RNAi-subjected genes (meiosis related).

Gene name	Predicted roles	Gene name	Predicted roles
rec-8	Meiotic cohesion	*syp-4*	Pairing and alignment
scc-3	Meiotic cohesion	*him-17*	DSB formation
smc-1	Meiotic cohesion	*mre-11*	DSB response
smc-3	Meiotic cohesion	*rad-50*	DSB formation
him-8	Pairing center	*xnp-1*	DSB formation
zim-1	Pairing center	*msh-4*	Strand exchange
zim-2	Pairing center	*him-14*	Strand exchange
zim-3	Pairing center	*msh-5*	Strand exchange
him-3	Pairing and alignment	*mlh-1*	Branch migration
htp-1	Pairing and alignment	*zyg-12*	Chromosome movement
htp-2	Pairing and alignment	*sun-1*	Chromosome movement
htp-3	Pairing and alignment	*plk-2*	Chromosome movement
syp-1	Pairing and alignment	*rfs-1*	Homologous recombination
syp-2	Pairing and alignment	*eme-1*	Resolution
zhp-3	Pairing and alignment	*R03H10.6*	*rpa-1* paralog
syp-3	Pairing and alignment		

Table 5. The list of putative TALT regions.

	Putative TALT			Upstream telomere		Downstream telomere	
Chr	Start	End	Size (bp)	Size (bp)	Strand	Size (bp)	Strand
I	432	613	181	telomere	-	270	-
I	833769	834058	289	336	+	336	+
I	834298	834425	127	336	+	56	+
I	939855	940046	191	99	+	81	+
I	940126	940699	573	81	+	177	+
I	940875	942559	1684	177	+	112	-
I	942670	943268	598	112	-	136	-
I	2368868	2370725	1857	177	+	161	-
I	11450517	11451202	685	112	-	131	+
I	11470232	11470917	685	131	-	118	+
I	11933575	11934656	1081	224	+	145	+
I	11948730	11949771	1041	187	+	124	+
I	13903332	13903754	422	115	-	194	-
I	13903947	13904269	322	194	-	195	+
II	1241828	1242092	264	46	-	79	-
II	2633596	2634812	1216	87	-	58	+
II	2846628	2846816	188	66	-	66	-
II	2846881	2847755	874	66	-	285	-
II	2978215	2979030	815	186	+	410	-
II	3014265	3015382	1117	244	-	95	-
II	4105552	4106574	1022	44	+	87	+
II	4152339	4153475	1136	98	+	194	-
II	4153475	4155194	1719	194	-	286	+
II	11928576	11930279	1703	256	+	147	-
II	11930425	11931451	1026	147	-	130	+
II	11931580	11932300	720	130	+	212	+
II	12169808	12169925	117	255	+	140	+
II	13974442	13975538	1096	89	+	61	+
II	14117099	14117221	122	94	-	64	-
II	14206600	14208350	1750	275	-	309	+
II	14819181	14819688	507	44	-	262	-
II	15253875	15254955	1080	107	-	673	+
III	1256340	1257447	1107	118	+	239	+
III	1264779	1265888	1109	116	+	239	+
III	1717653	1718814	1161	246	-	309	+
III	11784927	11785050	123	95	-	59	-
III	12130661	12131741	1080	159	-	496	+
III	12640910	12642422	1512	203	+	605	+
III	12944544	12945114	570	122	-	216	+
IV	1069499	1071380	1881	66	-	97	+
V	1230092	1230608	516	137	+	272	-
V	1318725	1318828	103	179	-	87	-
V	1480585	1481676	1091	187	-	255	-
V	1481930	1483211	1281	255	-	354	+
V	2708092	2709412	1320	667	-	305	-
V	16593232	16594124	892	220	+	104	+
V	18566072	18566408	336	420	-	267	-
V	18638399	18639697	1298	325	-	175	+
V	19229654	19231120	1466	584	+	451	+
V	19231570	19232597	1027	451	+	229	-
V	19722498	19723337	839	75	-	32	+
V	19827823	19829319	1496	223	-	352	+
V	20001786	20001928	142	75	+	105	+
V	20002032	20002324	292	105	+	558	+
X	1836510	1836859	349	208	-	531	-

Table 6. Strains having high copy of TALT2.

isotype	chrom	start	end	name	length	depth_of_coverage	genome_coverage	norm_coverage
JU1212	I	433	613	TALT_2	181	219.34	41.81	5.25
JU258	I	433	613	TALT_2	181	598.33	127.2	4.7
JU360	I	433	613	TALT_2	181	721.5	164.8	4.38
NIC1	I	433	613	TALT_2	181	235.56	68.39	3.44
NIC2	I	433	613	TALT_2	181	242.01	70.39	3.44
N2	I	433	613	TALT_2	181	485.15	145.59	3.33
NIC255	I	433	613	TALT_2	181	104.02	31.61	3.29
EG4347	I	433	613	TALT_2	181	216.87	67.72	3.2
WN2002	I	433	613	TALT_2	181	154.74	48.34	3.2
JU1896	I	433	613	TALT_2	181	221.13	72.55	3.05
JU1568	I	433	613	TALT_2	181	200.55	67.27	2.98
JU1395	I	433	613	TALT_2	181	61.62	21.11	2.92
CB4858	I	433	613	TALT_2	181	164.04	56.52	2.9
NIC251	I	433	613	TALT_2	181	94.35	36.47	2.59
NIC259	I	433	613	TALT_2	181	68.91	27.14	2.54
JU1580	I	433	613	TALT_2	181	382.13	151.06	2.53
JU1586	I	433	613	TALT_2	181	315.13	125.49	2.51
JU394	I	433	613	TALT_2	181	241.69	97.81	2.47
AB1	I	433	613	TALT_2	181	282.02	116.65	2.42
NIC260	I	433	613	TALT_2	181	92.66	38.67	2.4
JU561	I	433	613	TALT_2	181	106.33	44.71	2.38
JU1242	I	433	613	TALT_2	181	137.93	60.74	2.27
PB303	I	433	613	TALT_2	181	147.34	65.23	2.26
CB4851	I	433	613	TALT_2	181	266.48	122.47	2.18
LSJ1	I	433	613	TALT_2	181	238.96	111.27	2.15
JU2316	I	433	613	TALT_2	181	50.01	23.72	2.11
ED3012	I	433	613	TALT_2	181	124.39	59.83	2.08
JU2526	I	433	613	TALT_2	181	66.72	32.57	2.05
QG2075	I	433	613	TALT_2	181	62.43	30.4	2.05
ED3073	I	433	613	TALT_2	181	290.96	146.31	1.99
ED3017	I	433	613	TALT_2	181	247	129.75	1.9
JU2519	I	433	613	TALT_2	181	45.81	24.1	1.9
CB4932	I	433	613	TALT_2	181	219.5	116.82	1.88
NIC207	I	433	613	TALT_2	181	83.71	45.26	1.85
JU1200	I	433	613	TALT_2	181	273.19	150.44	1.82
JU1213	I	433	613	TALT_2	181	110.37	61.91	1.78
JU642	I	433	613	TALT_2	181	171.46	103.64	1.65
PX179	I	433	613	TALT_2	181	144.17	87.99	1.64
CB4852	I	433	613	TALT_2	181	482.44	297.43	1.62
JU440	I	433	613	TALT_2	181	102.95	68.12	1.51

References

Andersen, E.C., Gerke, J.P., Shapiro, J.A., Crissman, J.R., Ghosh, R., Bloom, J.S., Félix, M.-A., and Kruglyak, L. (2012). Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nature genetics *44*, 285-290.

Baird, D., and Royle, N. (1997). Sequences from higher primates orthologous to the human Xp/Yp telomere junction region reveal gross rearrangements and high levels of divergence. Human molecular genetics *6*, 2291-2299.

Blackburn, E.H. (1991). Structure and function of telomeres. Nature *350*, 569-573.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, btu170.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A., and Reddel, R.R. (1997). Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nature medicine *3*, 1271-1274.

Campisi, J. (2001). Cellular senescence as a tumor-suppressor mechanism. Trends in cell biology *11*, S27-S31.

Cesare, A.J., Kaul, Z., Cohen, S.B., Napier, C.E., Pickett, H.A., Neumann, A.A., and Reddel, R.R. (2009). Spontaneous occurrence of telomeric DNA damage response in the absence of chromosome fusions. Nature structural & molecular biology *16*, 1244-1251.

Cesare, A.J., and Reddel, R.R. (2010). Alternative lengthening of telomeres: models, mechanisms and implications. Nature reviews genetics *11*, 319-330.

Cheng, C., Shtessel, L., Brady, M.M., and Ahmed, S. (2012). Caenorhabditis elegans POT-2 telomere protein represses a mode of alternative lengthening of telomeres with normal telomere lengths. Proceedings of the National Academy of Sciences *109*, 7805-7810.

Cho, Nam W., Dilley, Robert L., Lampson, Michael A., and Greenberg, Roger A. (2014). Interchromosomal Homology Searches Drive Directional ALT Telomere Movement and Synapsis. Cell *159*, 108-121.

Conomos, D., Reddel, R.R., and Pickett, H.A. (2014). NuRD–ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination. Nature structural & molecular biology *21*, 760-770.

Conomos, D., Stutz, M.D., Hills, M., Neumann, A.A., Bryan, T.M., Reddel, R.R., and Pickett, H.A. (2012). Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells. The Journal of cell biology *199*, 893-906.

Curcio, M.J., and Belfort, M. (2007). The beginning of the end: links between ancient retroelements and modern telomerases. Proceedings of the National Academy of Sciences *104*, 9107-9108.

De Lange, T. (2004). T-loops and the origin of telomeres. Nature reviews Molecular cell biology *5*, 323-329.

De Lange, T. (2015). A loopy view of telomere evolution. Frontiers in genetics *6*.

Dvořáčková, M., Fojtová, M., and Fajkus, J. (2015). Chromatin dynamics of plant

telomeres and ribosomal genes. The Plant Journal.

Fajkus, J., Sýkorová, E., and Leitch, A.R. (2005). Telomeres in evolution and evolution of telomeres. Chromosome Research *13*, 469-479.

Fanti, L., Giovinazzo, G., Berloco, M., and Pimpinelli, S. (1998). The heterochromatin protein 1 prevents telomere fusions in Drosophila. Molecular cell *2*, 527-538.

Fasching, C.L., Bower, K., and Reddel, R.R. (2005). Telomerase-independent telomere length maintenance in the absence of alternative lengthening of telomeres–associated promyelocytic leukemia bodies. Cancer Research *65*, 2722-2729.

Fulcher, N., Derboven, E., Valuchova, S., and Riha, K. (2014). If the cap fits, wear it: an overview of telomeric structures over evolution. Cellular and Molecular Life Sciences *71*, 847-865.

Garavís, M., González, C., and Villasante, A. (2013). On the Origin of the Eukaryotic Chromosome: The Role of Noncanonical DNA Structures in Telomere Evolution. Genome biology and evolution *5*, 1142-1150.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell *43*, 405-413.

Heaphy, C.M., de Wilde, R.F., Jiao, Y., Klein, A.P., Edil, B.H., Shi, C., Bettegowda, C., Rodriguez, F.J., Eberhart, C.G., and Hebbar, S. (2011a). Altered telomeres in tumors with ATRX and DAXX mutations. Science *333*, 425-425.

Heaphy, C.M., Subhawong, A.P., Hong, S.-M., Goggins, M.G., Montgomery, E.A., Gabrielson, E., Netto, G.J., Epstein, J.I., Lotan, T.L., and Westra, W.H. (2011b). Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. The American journal of pathology *179*, 1608-1615.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Henson, J.D., Cao, Y., Huschtscha, L.I., Chang, A.C., Au, A.Y., Pickett, H.A., and Reddel, R.R. (2009). DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. Nature biotechnology *27*, 1181-1185.

Henson, J.D., Hannay, J.A., McCarthy, S.W., Royds, J.A., Yeager, T.R., Robinson, R.A., Wharton, S.B., Jellinek, D.A., Arbuckle, S.M., and Yoo, J. (2005). A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas. Clinical Cancer Research *11*, 217-225.

Jain, D., Hebden, A.K., Nakamura, T.M., Miller, K.M., and Cooper, J.P. (2010). HAATI survivors replace canonical telomeres with blocks of generic heterochromatin. Nature *467*, 223-227.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L.

(2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Lackner, D.H., Raices, M., Maruyama, H., Haggblom, C., and Karlseder, J. (2012). Organismal propagation in the absence of a functional telomerase pathway in Caenorhabditis elegans. The EMBO journal *31*, 2024-2033.

Lee, M., Hills, M., Conomos, D., Stutz, M.D., Dagg, R.A., Lau, L.M., Reddel, R.R., and Pickett, H.A. (2014). Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes. Nucleic acids research *42*, 1733-1746.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Linardopoulou, E.V., Williams, E.M., Fan, Y., Friedman, C., Young, J.M., and Trask, B.J. (2005). Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature *437*, 94-100.

Lundblad, V., and Blackburn, E.H. (1993). An alternative pathway for yeast telomere maintenance rescues *est1*− senescence. Cell *73*, 347-360.

Malik, H.S., Burke, W.D., and Eickbush, T.H. (2000). Putative telomerase catalytic subunits from Giardia lamblia and Caenorhabditis elegans. Gene *251*,

101-108.

Marciniak, R.A., Cavazos, D., Montellano, R., Chen, Q., Guarente, L., and Johnson, F.B. (2005). A novel telomere structure in a human alternative lengthening of telomeres cell line. Cancer research *65*, 2730-2737.

Mason, J.M., and Biessmann, H. (1995). The unusual telomeres of Drosophila. Trends in genetics *11*, 58-62.

Mason, J.M., Randall, T.A., and Capkova Frydrychova, R. (2015). Telomerase lost? Chromosoma.

Mason, J.M., Reddy, H.M., and Frydrychova, R.C. (2011). Telomere maintenance in organisms without telomerase (INTECH Open Access Publisher).

McClintock, B. (1942). The fusion of broken ends of chromosomes following nuclear fusion. Proceedings of the national academy of sciences of the United States of America *28*, 458.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., and Daly, M. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research *20*, 1297-1303.

Mefford, H.C., and Trask, B.J. (2002). The complex structure and dynamic evolution of human subtelomeres. Nature Reviews Genetics *3*, 91-102.

Meier, B., Clejan, I., Liu, Y., Lowden, M., Gartner, A., Hodgkin, J., and Ahmed, S. (2006). trt-1 is the Caenorhabditis elegans catalytic subunit of telomerase. PLoS

genetics *2*, e18.

Moore, J.K., and Haber, J.E. (1996). Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature *383*, 644-646.

Muller, H., and Herskowitz, I.H. (1954). Concerning the healing of chromosome ends produced by breakage in Drosophila melanogaster. American Naturalist, 177-208.

Nakamura, T.M., Cooper, J.P., and Cech, T.R. (1998). Two modes of survival of fission yeast without telomerase. Science *282*, 493-496.

Nosek, J., Kosa, P., and Tomaska, L. (2006). On the origin of telomeres: a glimpse at the pre-telomerase world. Bioessays *28*, 182-190.

O'Sullivan, R.J., and Almouzni, G. (2014). Assembly of telomeric chromatin to create ALTernative endings. Trends in cell biology.

Olovnikov, A.M. (1973). A theory of marginotomy: the incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. Journal of theoretical biology *41*, 181-190.

Pich, U., and Schubert, I. (1998). Terminal heterochromatin and alternative telometric sequences in Allium cepa. Chromosome Research *6*, 315-322.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Riethman, H., Ambrosini, A., and Paul, S. (2005). Human subtelomere structure and variation. Chromosome Research *13*, 505-515.

Schwartzentruber, J., Korshunov, A., Liu, X.-Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.-A.K., and Tönjes, M. (2012). Driver mutations in histone H3. 3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226-231.

Seo, B., Kim, C., Hills, M., Sung, S., Kim, H., Kim, E., Lim, D.S., Oh, H.-S., Choi, R.M.J., Chun, J., *et al.* (2015). Telomere maintenance through recruitment of internal genomic regions. Nat Commun *6*.

Shay, J., and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. European journal of cancer *33*, 787-791.

Shay, J.W., and Wright, W.E. (2010). Telomeres and telomerase in normal and cancer stem cells. FEBS letters *584*, 3819-3825.

Shtessel, L., Lowden, M.R., Cheng, C., Simon, M., Wang, K., and Ahmed, S. (2013). Caenorhabditis elegans POT-1 and POT-2 repress telomere maintenance pathways. G3: Genes| Genomes| Genetics *3*, 305-313.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., and Lander, E.S. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.

Sykorova, E., Lim, K.Y., Chase, M.W., Knapp, S., Leitch, I.J., Leitch, A.R., and Fajkus, J. (2003). The absence of Arabidopsis-type telomeres in Cestrum and

closely related genera Vestia and Sessea (Solanaceae): first evidence from eudicots. The Plant Journal *34*, 283-291.

Teng, S.-C., Kim, B., and Gabriel, A. (1996). Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. Nature *383*, 641-644.

Teng, S.-C., and Zakian, V.A. (1999). Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in Saccharomyces cerevisiae. Molecular and cellular biology *19*, 8083-8093.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*, 562-578.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology *28*, 511-515.

Varley, H., Pickett, H.A., Foxon, J.L., Reddel, R.R., and Royle, N.J. (2002). Molecular characterization of inter-telomere and intra-telomere mutations in human ALT cells. Nature genetics *30*, 301-305.

국문초록

대안적 텔로미어 유지기작의 새로운 DNA 복제단위 연구

김천아

서울대학교 생명과학부

DNA 중합효소는 염색체 말단 부분인 텔로미어를 복제할 수 없다. 이를 ‘말단 복제 문제’라고 부른다. 만약 세포가 이 문제를 해결하지 못하면 염색체는 세포분열을 거듭할 때마다 짧아지게 된다. 염색체를 온전히 유지하기 위해 개체는 텔로미어 유지 기작을 개발했다. 고대 진핵세포에서 특별한 역전사효소인 텔로머레이즈가 진화했다. 그런데 흥미롭게도 다양한 종이 진화하는 동안 텔로머레이즈 유전자 결손이 여러차례 발생하였다. 텔로머레이즈 결손에 대처하기 위해 다양한 종은 텔로머레이즈 이외의 다른 텔로미어 유지 기작을 채택하였다. 대안적 텔로미어 유지기작 (ALT) 은 암세포 연구를 통해 잘 알려진 텔로머레이즈 비-의존적인 텔로미어 유지기작이다. 그러나 지금까지 수행된 ALT 연구는 주로 세포 수준에 국한되어 있었고, 개체 수준에서 나타나는 변화양상에 대한 연구는 부족한 상황이다. 그래서 나는 텔로머레이즈 유전자가 결손된 동물 모델인 예쁜꼬마선충을 이용하여 텔로머레이즈 비-의

존적인 텔로미어 유지기작을 개체 수준에서 관찰해보고자 하였다. 본 연구에서 나는 예쁜꼬마선충에서 텔로머레이즈 비-의존적으로 텔로미어 길이를 늘릴 수 있는 개체를 찾아냈다. 흥미롭게도 이들의 텔로미어는 텔로머레이즈가 염색체 말단에 삽입하는 전형적인 텔로미어 반복 서열이 아닌 다른 형태의 긴 DNA 단위의 반복으로 이루어져 있었다. 우리는 이 복제 단위를 TALT 라고 명명하였다. TALT 서열은 전형적인 텔로미어 반복서열과 연관이 없는 특이한 긴 DNA를 텔로미어와 유사한 반복서열이 감싸고 있는 형태다. 염색체 내부에 있던 TALT는 텔로머레이즈 결손이 일어나기 전에 염색체 끝 부분으로 이동하여 TALT 저장소를 만들어 두었다가 텔로미어에 위기가 발생하였을 때 전체 텔로미어로 복제된다. 이렇게 텔로미어 서열이 TALT로 변화한 상황은 텔로미어 진화의 단면을 제시해준다. 또한, 텔로머레이즈 결손이 일어나기 전 TALT가 염색체 말단으로 옮겨가는 현상은 서브 텔로미어 (텔로미어 인접 부위) 진화의 새로운 분자적 기전을 제안한다.

주요어: 예쁜꼬마선충, 유전학, 텔로미어, 텔로미어 유지 기작, 대안적 텔로미어 유지기작, 텔로머레이즈 비-의존적인 동물모델, TALT

학번: 2009-20328